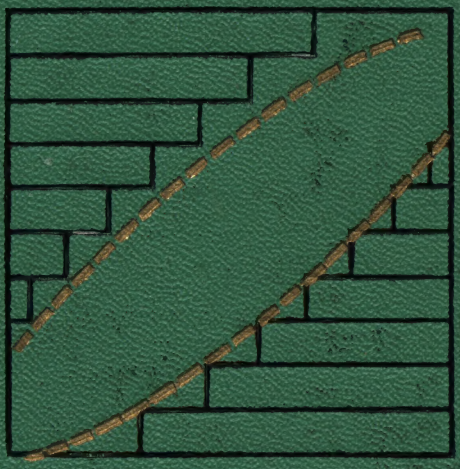
М. Дж. Кендалл
А. Стьюарт

СТАТИСТИЧЕСКИЕ ВЫВОДЫ И СВЯЗИ





THE ADVANCED THEORY OF STATISTICS

MAURICE G. KENDALL, M.A., Sc.D.

Managing Director, C-E-I-R Ltd . Formerly Professor of Statistics in the University of London President of the Royal Statistical Society, 1960–1962

and

ALAN STUART, D.Sc. (ECON.)

Professor of Statistics in the University of London

IN THREE VOLUMES

VOLUME 2

INFERENCE AND RELATIONSHIP

SECOND EDITION



CHARLES GRIFFIN & COMPANY LIMITED

LONDON

ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1973

© Перевод на русский язык, «Наука», 1973.

К $\frac{0223\text{-}1852}{042\,(02)\text{-}73}$ 45-73

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ РЕДАКТОРА ПЕРЕВОДА 7
ПРЕДИСЛОВИЕ К ПЕРВОМУ ИЗДАНИЮ 11
ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ 12

ГЛАВЫ

17. ОЦЕНИВАНИЕ 13
18. ОЦЕНИВАНИЕ: МЕТОД МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ . 58
19. ОЦЕНИВАНИЕ: МЕТОД НАИМЕНЬШИХ КВАДРАТОВ И ДРУГИЕ МЕТОДЫ 109
20. ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ: ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ 139
21. ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ: ФИДУЦИАЛЬНЫЕ ИНТЕРВАЛЫ 183
22. ПРОВЕРКА ГИПОТЕЗ: ПРОСТЫЕ ГИПОТЕЗЫ 219
23. ПРОВЕРКА ГИПОТЕЗ: СЛОЖНЫЕ ГИПОТЕЗЫ 249
24. КРИТЕРИИ ОТНОШЕНИЯ ПРАВДОПОДОБИЯ И ОБЩАЯ ЛИНЕЙНАЯ ГИПОТЕЗА 301
25. СРАВНЕНИЕ КРИТЕРИЕВ 351
26. СТАТИСТИЧЕСКАЯ ЗАВИСИМОСТЬ: ЛИНЕЙНАЯ РЕГРЕССИЯ И КОРРЕЛЯЦИЯ 373
27. ЧАСТНАЯ И МНОЖЕСТВЕННАЯ КОРРЕЛЯЦИЯ 423
28. ОБЩАЯ ТЕОРИЯ РЕГРЕССИИ 461
29. ФУНКЦИОНАЛЬНАЯ И СТРУКТУРНАЯ ЗАВИСИМОСТЬ 500
30. КРИТЕРИИ СОГЛАСИЯ 558
31. УСТОЙЧИВЫЕ И СВОБОДНЫЕ ОТ РАСПРЕДЕЛЕНИЯ ПРОЦЕДУРЫ 622
32. НЕКОТОРЫЕ ПРИМЕНЕНИЯ ПОРЯДКОВЫХ СТАТИСТИК 687
33. КАТЕГОРИЗОВАННЫЕ ДАННЫЕ 719
34. ПОСЛЕДОВАТЕЛЬНЫЕ МЕТОДЫ 791

Приложение

ТАБЛИЦЫ

1. ФУНКЦИЯ ПЛОТНОСТИ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ . . . 833
2. ФУНКЦИЯ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ 833
3. КВАНТИЛИ χ^2-РАСПРЕДЕЛЕНИЯ 835
4а. ФУНКЦИЯ χ^2-РАСПРЕДЕЛЕНИЯ С ОДНОЙ СТЕПЕНЬЮ СВОБОДЫ ДЛЯ $0 \leqslant \chi^2 \leqslant 1$ 836

46. ФУНКЦИЯ χ^2-РАСПРЕДЕЛЕНИЯ С ОДНОЙ СТЕПЕНЬЮ СВОБОДЫ ДЛЯ $1 \leqslant \chi^2 \leqslant 10$ 836
5. КВАНТИЛИ t-РАСПРЕДЕЛЕНИЯ 837
6. 5%-НЫЕ ТОЧКИ z-РАСПРЕДЕЛЕНИЯ 839
7. 5%-НЫЕ ТОЧКИ F-РАСПРЕДЕЛЕНИЯ 840
8. 1%-НЫЕ ТОЧКИ z-РАСПРЕДЕЛЕНИЯ 841
9. 1%-НЫЕ ТОЧКИ F-РАСПРЕДЕЛЕНИЯ 842

ЦИТИРОВАННАЯ ЛИТЕРАТУРА 844

УКАЗАТЕЛЬ 878

ПРЕДИСЛОВИЕ РЕДАКТОРА ПЕРЕВОДА

Публикуемая сейчас книга является переводом второго тома трехтомного труда М. Кендалла и А. Стьюарта «The Advanced Theory of Siatistics». Первый том появился на русском языке в 1966 г. под названием «Теория распределений». Первые издания всего труда, написанные одним Кендаллом, выходили в двух томах с 1943 по 1952 г. Это была наиболее полная в мировой литературе сводка получивших практические применения сведений и формул, сопровождаемая большим числом умело подобранных примеров. В 1958 г. начали выходить тома трехтомного издания, составленного совместно Кендаллом и Стьюартом. Это расширенное издание закончено в 1966 г. Появившийся в 1966 г. третий том тоже предполагается перевести на русский язык.

Несмотря на участие младшего из авторов (А. Стьюарта), трехтомное издание остается несколько старомодным (каким было двухтомное уже в момент своего появления). Читатели, знакомые с современным стилем изложения теории вероятностей и математической статистики, соединяющим общность и логическую отчетливость с простотой и краткостью, местами будут разочарованы. Но богатство конкретного материала и обилие показательных примеров искупают недостатки труда Кендалла и Стьюарта.

Терминология по возможности согласована с переводом первого тома. Небольшие погрешности исправлены без специальных на то указаний. В более ответственных случаях отклонения от подлинника оговорены в примечаниях. За примечания «от редактора» я несу

ответственность совместно с Д. М. Чибисовым — инициатором большинства из них.

Многие из упомянутых авторами таблиц имеются в книгах: Л. Н. Большев, Н. В. Смирнов, *Таблицы математической статистики*, М., изд-во «Наука», 1965; Д. Б. Оуэн, *Сборник статистических таблиц* (перев. с англ.), М., Изд-во Вычислительного центра АН СССР, 1966. В таких случаях сделаны примечания со ссылками на эти книги.

А. Н. Колмогоров

«Вы еще не рассказали мне, — заметила леди Наттэл, — чем занимается ваш жених».

«Он статистик», — ответила Лэймия с неприятным ощущением обороняющегося.

Леди Наттэл была, очевидно, застигнута врасплох. Ей не приходило в голову, что статистики вступают в нормальные общественные отношения. Этот род людей, как она предполагала, увековечен на некоторый второстепенный образ действий, наподобие мулов.

«Но, тетя Сара, это очень интересная профессия», — сказала с жаром Лэймия.

«Я не сомневаюсь в этом, — ответила ее тетя, которая, очевидно, очень сомневалась. — Выразить что-нибудь значительное в одних лишь цифрах, понятно, настолько невозможно, что должен существовать бесконечный простор для хорошо оплаченных советов относительно того, как это сделать. Но не думаете ли вы, что жизнь со статистиком была бы до некоторой степени, скажем, скучной?»

Лэймия молчала. Ей не хотелось говорить о поразительной глубине эмоциональных возможностей, которую она открыла под цифровой внешностью занятий Эдварда.

«Дело не в цифрах, — сказала она наконец, — а в том, что вы с ними делаете».

К. Мэндервил, Крушение Лэймии Гёрдлнек

(K. A. C. Manderville, The Undoing of Lamia Gurdleneck)

ПРЕДИСЛОВИЕ К ПЕРВОМУ ИЗДАНИЮ

Предлагаемая книга является вторым томом трактата о современной теории статистики. В нее входят теория оценивания и проверки гипотез, статистические связи, свободные от распределения методы и последовательный анализ. Третий, заключительный том будет содержать планирование и анализ выборочных обследований и экспериментов, включая дисперсионный анализ, а также теорию многомерного анализа и временные ряды.

Настоящий том имеет очень слабое сходство с первоначальным вторым томом «Современной теории» Кендалла. Его пришлось перестроить и переписать практически заново вследствие быстрого развития статистики в последние пятнадцать лет. Беглый взгляд на прилагаемый список литературы покажет, как много работ было опубликовано за этот период.

Мы старались сделать данный том, как и первый, самостоятельным в трех отношениях: он имеет свой список литературы, снабжен собственным указателем, и в нем повторяются необходимые таблицы, приведенные в приложении к первому тому. Нам не пришлось тратить слишком много места для обширной библиографии, поскольку Кендалл и Дойг (Kendall and Doig) отдельно опубликовали исчерпывающую Bibliography of Statistical Literature. Мы специально позаботились о том, чтобы снабдить изложение большим количеством упражнений: их в данном томе имеется около 400.

За разрешение воспроизвести некоторые из таблиц, помещенных в конце книги, мы весьма признательны проф. сэру Рональду Фишеру (sir Ronald Fisher), д-ру Фрэнку Иэйтсу (Frank Yates), м-рам Оливеру и Бойду (Oliver and Boyd) и издателям журнала Biometrika. М-р Э. В. Бёрк (E. V. Burke) из издательства Charles Griffin and Company Limited оказал нам неоценимую помощь при печатании книги. Мы также признательны м-ру К. Мэндервилу (K. A. C. Menderville) за разрешение процитировать из его неопубликованного рассказа отрывок, приведенный на странице 9.

Как всегда, мы будем благодарны за уведомления о каких-либо ошибках, опечатках или неясностях.

Лондон,
март, 1961

М. Дж. К.,
А. С.

ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ

Помимо устранения несущественных опечаток и ошибок, этот том был полностью просмотрен и исправлен для данного издания. В нескольких местах прогресс статистики сделал необходимым переписать заново страницу или две, хотя обычно требовалось изменение только отдельных абзацев или предложений. В процессе переработки было добавлено несколько сотен новых ссылок и, кроме того, значительное количество новых упражнений, в большинстве взятых из последних работ.

Мы благодарны читателям всего мира, которые помогли нам сделать эти улучшения.

Лондон,
январь, 1967

М. Дж. К.,
А. С.

ГЛАВА 17

ОЦЕНИВАНИЕ

Постановка задачи

17.1 В предшествующих главах нам в ряде случаев приходилось сталкиваться с задачей оценивания параметров генеральной совокупности по выборочным данным. Однако до сих пор обсуждение такого рода вопросов проводилось, по существу, на интуитивном уровне. Так, в теории больших выборок мы брали в качестве оценок среднего и моментов генеральной совокупности выборочное среднее и выборочные моменты, считая эти оценки вполне удовлетворительными.

Теперь мы приступаем к более детальному изучению этого предмета. В настоящей главе будут рассмотрены различные критерии, которым должна удовлетворять «хорошая» оценка, и изучен вопрос о существовании «наилучших» (в приемлемом смысле этого слова) оценок. Несколько следующих глав будут посвящены методам получения оценок с требуемыми свойствами.

17.2 Очевидно, что если выборка не случайна, а получена каким-то целенаправленным способом, относительно которого ничего определенного не известно, то о генеральной совокупности можно сказать очень немногое. Кое-какие выводы тривиального характера, конечно, возможны. Например, если мы берем десять реп из кучи в 100 реп и находим, что их общий вес равен десяти фунтам, то ясно, что средний вес одной репы в куче не меньше одной десятой фунта. Однако информация такого рода вряд ли может представлять ценность. Что касается оценивания на основе пристрастных выборок, то оно принадлежит скорее области субъективных суждений, нежели области точных и объективных понятий. Поэтому мы ограничимся только случайными выборками. Общая задача, в ее простейшей постановке, будет заключаться в оценивании параметра генеральной совокупности на основе информации, даваемой выборкой. Мы начнем с задачи оценивания одного параметра. Случай нескольких параметров будет рассмотрен позже.

17.3 Выясним прежде всего, что понимается под словом «оценивание». Предположим, что мы знаем, с точностью до значения неизвестного параметра θ, распределение генеральной совокупности. Пусть, кроме того, дана выборка из n наблюдений $x_1, \ldots, x_n$. Требуется определить, используя эти наблюдения, число, которое можно было бы взять в качестве значения θ, или интервал, о котором можно было бы утверждать, что он содержит это значение.

Будем называть статистикой любую функцию, зависящую только от наблюдений. Так как наблюдения являются случайными величинами, то статистики будут тоже случайными величинами. Поэтому, если какая-то из них используется для оценивания параметра θ, то для отдельных выборок могут получиться значения, сильно отличающиеся от истинного. Отсюда ясно, что нельзя найти оценку, которая принимала бы значения, близкие к θ, для всех возможных выборок. Мы должны ограничиться такой процедурой оценивания, которая давала бы хорошие результаты «в среднем» при многократном ее использовании или имела бы «большую вероятность успеха». Другими словами, надо иметь в виду, что методы оценивания порождают распределения значений оценок, и сравнивать их достоинства, исходя из свойств этих распределений.

17.4 Мы достигнем большей ясности, если будем различать правило оценивания (которое назовем оценочной функцией или просто оценкой) и значения, к которым это правило приводит в конкретных случаях*). Различие между ними такое же, как между функцией $f(x)$, определенной на некотором множестве значений переменной x, и отдельным значением, скажем $f(a)$, которое эта функция принимает при фиксированном значении x, равном a. Наша задача заключается в нахождении оценок, а не их частных значений. Мы не будем отвергать оценку, если она даст плохие результаты для отдельных выборок (в том смысле, что ее значения будут существенно отличаться от истинного значения), но мы ее отвергнем, если она будет давать плохие результаты очень часто, т. е. если распределение значений оценки будет сильно расходиться с истинным значением θ. Вообще, качество оценки будет определяться целиком на основании ее выборочного распределения.

17.5 В теории больших выборок мы часто брали в качестве оценки для θ статистику t, вычисляемую по выборке точно таким же путем, каким θ вычислялось по генеральной совокупности. Например, выборочное среднее считалось оценкой для генерального среднего. Попробуем обосновать эту процедуру.

---

*) Соответствующие английские термины — estimator (оценка) и estimate (значение оценки). (*Прим. ред.*)

Рассмотрим случай генеральной совокупности вида

$$dF(x) = (2\pi)^{-1/2} \exp\left\{-\frac{1}{2}(x-\theta)^2\right\} dx, \quad -\infty \leqslant x \leqslant \infty. \quad (17.1)$$

В качестве оценки генерального среднего θ возьмем выборочное среднее

$$t = \frac{1}{n} \sum_{i=1}^{n} x_i. \quad (17.2)$$

Распределение t имеет вид (пример 11.12)

$$dF(t) = \{n/2\pi\}^{1/2} \exp\left\{-\frac{1}{2} n (t-\theta)^2\right\} dt, \quad (17.3)$$

т. е. t распределено нормально со средним θ и дисперсией $1/n$.

Отметим две особенности этого распределения: (а) его среднее, а также мода и медиана равны истинному значению θ; (б) при возрастании n рассеяние значений t относительно θ становится меньше, и вероятность того, что данное значение t отличается более чем на некоторую фиксированную величину от θ, уменьшается. Можно сказать, что точность оценки увеличивается с ростом n.

17.6 Вообще, нетрудно понять, что выражение «точность, увеличивающаяся с ростом n» имеет определенный смысл в том случае, когда дисперсия выборочного распределения t убывает с убыванием $1/n$, а центральное значение либо совпадает с θ, либо отличается от него на величину, которая также убывает с убыванием $1/n$. Большинство оценок, с которыми обычно имеют дело, принадлежит к этому типу. Но есть и исключения. Рассмотрим, например, выборку из генеральной совокупности, имеющей распределение Коши:

$$dF(x) = \frac{1}{\pi} \frac{dx}{\{1 + (x-\theta)^2\}}, \quad -\infty \leqslant x \leqslant \infty. \quad (17.4)$$

Если мы возьмем в качестве оценки для θ выборочное среднее t, то распределение t будет иметь вид

$$dF(t) = \frac{1}{\pi} \frac{dt}{\{1 + (t-\theta)^2\}} \quad (17.5)$$

(см. пример 11.1). В этом случае распределение t совпадает с распределением одного наблюдения и точность не увеличивается с ростом n.

Состоятельность

17.7 Увеличение точности с увеличением объема выборки является, конечно, очень желательным свойством оценки. В случае, когда дисперсия выборочного распределения оценки убывает с ростом n, для возрастания точности необходимо еще,

чтобы центральное значение этого распределения стремилось к θ, иначе оценка будет принимать значения, систематически отличающиеся от истинного. Поэтому первый критерий качества оценки может быть сформулирован следующим образом.

Оценка t_n, вычисленная на основании выборки объема n, называется состоятельной оценкой для θ, если для любых сколь угодно малых положительных ε и η существует N такое, что вероятность неравенства

$$|t_n - \theta| < \varepsilon \tag{17.6}$$

больше $1 - \eta$ для всех $n > N$, или, в обозначениях теории вероятностей,

$$\mathbf{P}\{|t_n - \theta| < \varepsilon\} > 1 - \eta, \quad n > N. \tag{17.7}$$

Приведенное определение аналогично обычному определению сходимости числовой последовательности. Для любого наперед заданного числа ε можно найти такое достаточное большое число, что при любом объеме выборки, превышающем это число, вероятность того, что t отличается от истинного значения больше чем на ε, будет сколь угодно близка к нулю. Оценка t_n называется в этом случае *сходящейся по вероятности* или *стохастически* к θ. Таким образом, оценка t будет состоятельной оценкой для θ, если она сходится к θ по вероятности.

Пример 17.1

Выборочное среднее является состоятельной оценкой параметра θ генеральной совокупности (17.1). Это было установлено с помощью общих рассуждений, формальное же доказательство выглядело бы следующим образом.

Пусть ε задано. Согласно (17.3) $(t - \theta)\, n^{1/2}$ распределено нормально с нулевым средним и единичной дисперсией. Следовательно, вероятность события $|(t - \theta)\, n^{1/2}| \leqslant \varepsilon n^{1/2}$ равна значению нормального интеграла в пределах $\pm \varepsilon n^{1/2}$. Для любого положительного η всегда можно выбрать такое достаточно большое n, что это значение будет больше $1 - \eta$, причем неравенство будет выполняться и для бóльших n. Таким образом, число N, требуемое в определении сходимости по вероятности, может быть найдено, и, следовательно, условие (17.7) будет выполнено.

Пример 17.2

Пусть имеется статистика t_n со средним значением, отличающимся от θ членами порядка n^{-1}, и с дисперсией v_n порядка n^{-1}, и пусть распределение $(t_n - \theta)/v_n^{1/2}$ стремится к нормальному при возрастании n. Тогда, как и в примере 17.1, $(t_n - \theta)$ будет

стремиться по вероятности к нулю и t_n будет состоятельной оценкой для θ. Этот случай охватывает большинство статистик, встречающихся на практике.

Результат остается справедливым даже при неизвестном предельном распределении. Это непосредственно следует из неравенства Бьенэме — Чебышева (3.94). Действительно, если

$$\mathsf{M}(t_n) = \theta + k_n \quad \text{и} \quad \mathsf{D}t_n = v_n,$$

где

$$\lim_{n\to\infty} k_n = \lim_{n\to\infty} v_n = 0,$$

то мы имеем

$$\mathsf{P}\{|t_n - (\theta + k_n)| < \varepsilon\} \geqslant 1 - \frac{v_n}{\varepsilon^2} \xrightarrow[n\to\infty]{} 1,$$

т. е. (17.7) будет выполнено.

Несмещенные оценки

17.8 Свойство состоятельности является предельным свойством, т. е. относится к поведению оценки при стремлении объема выборки к бесконечности и не налагает никаких ограничений на поведение оценки при конечных n. Если существует одна состоятельная оценка t_n, то можно построить бесконечно много других. Например, при фиксированных a и b оценка

$$\frac{n-a}{n-b}\, t_n$$

будет также состоятельной. Мы знаем, что в некоторых случаях состоятельной оценкой генерального среднего является выборочное среднее

$$\bar{x} = \sum x_j/n.$$

Но статистика

$$\bar{x}' = \sum x_j/(n-1)$$

тоже будет состоятельной оценкой генерального среднего. Почему же мы одну предпочитаем другой? Интуитивно кажется абсурдным делить сумму n количеств на какое-либо число, отличное от n. Однако мы скоро увидим, что интуиция не очень надежный советчик в этом вопросе. Так, имеются основания для предпочтения статистики

$$\frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})^2$$

статистике

$$\frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{x})^2$$

в качестве оценки генеральной дисперсии, несмотря на то что последняя является выборочной дисперсией.

17.9 Рассмотрим выборочное распределение некоторой оценки t. Если оценка состоятельна, то центральное значение этого распределения при больших объемах выборки должно быть близко к θ. Можно дополнительно потребовать, чтобы это центральное значение было равно θ для всех выборок, а не только для больших.

Если требовать, чтобы для всех n среднее значение t было равно θ, т. е.

$$\mathsf{M}(t) = \theta,$$

то мы определим так называемую несмещенную оценку. Этот термин, подобно многим в статистике, нельзя назвать удачным. Нет никаких оснований, кроме удобства, выделять арифметическое среднее среди других мер расположения как критерий несмещенности. С таким же успехом мы могли бы выбрать для определения несмещенной оценки медиану распределения t или его моду. Среднее выбрано, как всегда, из-за его хороших математических свойств. Это вполне разумно, но надо помнить, что этому термину не следует придавать каких-либо оттенков не технического характера *).

Пример 17.3

Равенство

$$\mathsf{M}\left\{\frac{1}{n}\sum x\right\} = \frac{1}{n}\sum \mathsf{M}(x) = \mu_1'$$

показывает, что выборочное среднее является несмещенной оценкой генерального среднего всегда, когда последнее существует. Однако выборочная дисперсия не будет несмещенной оценкой генеральной дисперсии. Действительно,

$$\mathsf{M}\left\{\sum (x_i - \bar{x})^2\right\} = \mathsf{M}\left\{\sum \left[x_i - \sum_j x_j/n\right]^2\right\} =$$

$$= \mathsf{M}\left\{\frac{n-1}{n}\sum_i x_i^2 - \frac{1}{n}\sum_{j\neq k}\sum x_j x_k\right\} =$$

$$= (n-1)\,\mu_2' - (n-1)\,\mu_1'^2 = (n-1)\,\mu_2,$$

*) Несколько более подробное обсуждение вопроса о том, когда желательна несмещенность оценки имеется в книге: С. Р. Рао, *Линейные статистические методы и их применения*, М., «Наука», 1968, стр. 270—272. *(Прим. ред.)*

т. е. $\frac{1}{n}\sum (x-\bar{x})^2$ имеет среднее значение $\frac{n-1}{n}\mu_2$. Несмещенной оценкой будет

$$\frac{1}{n-1}\sum (x-\bar{x})^2,$$

поэтому ее обычно предпочитают выборочной дисперсии.

Из сказанного видно, что состоятельная оценка не обязательно является несмещенной. Нам встречалась также несмещенная оценка, которая не была состоятельной (пример 14.5). Таким образом, ни одно из этих свойств не влечет другое. Однако состоятельная оценка, асимптотическое распределение которой имеет конечное среднее, должна быть асимптотически несмещенной.

Иногда несмещенной оценки вообще не существует (см. упражнение 17.12). Если такая оценка и существует, то может случиться, что в отдельных случаях, и даже всегда, она будет давать бессмысленные результаты. Например, при оценивании параметра θ, $0 \leqslant \theta \leqslant 1$, никакая статистика, распределение которой сосредоточено на интервале $(0, 1)$, не будет несмещенной, так как для $\theta = 0$ ее математическое ожидание (за исключением тривиальных случаев вырождения) превосходит θ. Один важный пример такого рода будет рассмотрен в **27.34**. Упражнение 17.26 содержит принадлежащий Э. Леману пример несмещенной оценки, которая всегда приводит к бессмысленным результатам.

Поправки на смещение

17.10 Если имеется состоятельная, но смещенная оценка и желательно исключить смещение, то можно вычислить математическое ожидание оценки и попытаться найти нужную поправку, подобно тому как было сделано в примере 17.3. Однако это математическое ожидание иногда бывает довольно сложной функцией параметра θ, который нужно оценить, и тогда не очень понятно, какова должна быть поправка. Остроумный метод, позволяющий преодолевать эту трудность в широком классе случаев, предложил Кэнуй (1956).

Пусть (смещенная) оценка t_n, вычисленная на основе n наблюдений, является функцией выборочных k-статистик k_i (глава 12). Пусть, кроме того, существуют все семиинварианты $\varkappa_i$ генеральной совокупности (тогда, как известно, k_i будут несмещенными оценками для $\varkappa_i$). Разложение t_n в ряд Тейлора относительно θ дает

$$t_n-\theta = \sum (k_i-\varkappa_i)\left(\frac{\partial t_n}{\partial k_i}\right) + \frac{1}{2}\sum\sum (k_i-\varkappa_i)(k_j-\varkappa_j)\left(\frac{\partial^2 t_n}{\partial k_i\,\partial k_j}\right) + \ldots \tag{17.8}$$

(производные взяты в точках $k_i = \varkappa_i$). Беря затем математические ожидания от обеих частей равенства (17.8) и учитывая тот факт, что моменты k-статистик являются степенными рядами по степеням $1/n$ (12.13), мы получим

$$\mathbf{M}(t_n) - \theta = \sum_{r=1}^{\infty} \frac{a_r}{n^r}. \tag{17.9}$$

Отметим, что, даже если t_n не будет функцией k-статистик, (17.9) может иногда выполняться (см. упражнение 17.13). Пусть теперь $\bar{t}_{n-1}$ обозначает статистику, полученную осреднением статистик t_{n-1}, вычисленных для всех n подмножеств из $n-1$ наблюдений. Если рассмотреть новую статистику

$$t'_n = nt_n - (n-1)\bar{t}_{n-1}, \tag{17.10}$$

то из (17.9) немедленно будет следовать

$$\mathbf{M}(t'_n) - \theta = a_2\left(\frac{1}{n} - \frac{1}{n-1}\right) + a_3\left(\frac{1}{n^2} - \frac{1}{(n-1)^2}\right) + \ldots = -\frac{a_2}{n^2} + O(n^{-3}). \tag{17.11}$$

Последнее выражение показывает, что смещение t'_n имеет порядок $1/n^2$. Аналогичные вычисления показывают, что статистика

$$t''_n = \{n^2 t'_n - (n-1)^2 \bar{t}'_{n-1}\}/\{n^2 - (n-1)^2\}$$

будет уже иметь смещение порядка $1/n^3$, и т. д. Таким образом, предложенный метод позволяет исключать смещение с любой желаемой точностью.

В упражнении 17.18 требуется показать, что $\mathbf{D}t'_n = \mathbf{D}t_n$ с точностью до членов порядка $1/n$, т. е. исключение смещения первого порядка не увеличивает дисперсии. Это, однако, может быть неверно для поправок более высокого порядка.

В статье Миллера (1964) рассматривается вопрос об асимптотической нормальности статистики t'_n и об оценивании ее дисперсии.

Пример 17.4

Найти несмещенную оценку для θ^2, где θ — параметр биномиального распределения

$$\mathbf{P}\{x = r\} = \binom{n}{r}\theta^r(1-\theta)^{n-r}, \quad r = 0, 1, 2, \ldots, n.$$

Интуитивной оценкой будет

$$t_n = \left(\frac{r}{n}\right)^2,$$

поскольку r/n — несмещенная оценка для θ. Статистика t_{n-1} может принимать только два значения:

$$\left(\frac{r-1}{n-1}\right)^2 \quad \text{или} \quad \left(\frac{r}{n-1}\right)^2,$$

в соответствии с тем, что исключено из выборки: «успех» или «неудача». Следовательно,

$$\bar{t}_{n-1} = \frac{1}{n}\left\{r\left(\frac{r-1}{n-1}\right)^2 + (n-r)\left(\frac{r}{n-1}\right)^2\right\} = \frac{r^2(n-2)+r}{n(n-1)^2},$$

поэтому (17.10) примет вид

$$t'_n = nt_n - (n-1)\bar{t}_{n-1} = \frac{r^2}{n} - \frac{r^2(n-2)+r}{n(n-1)} = \frac{r(r-1)}{n(n-1)}. \quad (17.12)$$

Полученная оценка оказывается точной несмещенной оценкой для θ^2.

17.11 Вообще может существовать более одной состоятельной оценки параметра, даже если ограничиться только несмещенными оценками. Рассмотрим еще раз задачу оценивания среднего нормальной совокупности с дисперсией σ^2. Выборочное среднее является состоятельной и несмещенной оценкой. Покажем, что то же самое верно и для медианы.

Из соображений симметрии следует, что медиана является несмещенной оценкой генерального среднего, которое, конечно, совпадает с медианой генеральной совокупности. Для больших n распределение медианы стремится к нормальному распределению (см. 14.12):

$$dF(x) \propto \exp\{-2nf_1^2(x-\theta)^2\}\,dx \quad (17.13)$$

(f_1 — ордината медианы генеральной совокупности *), равная $(2\pi\sigma^2)^{-1/2}$ в нашем случае). Из (17.13) видно, что дисперсия выборочной медианы равна $\pi\sigma^2/(2n)$ и стремится к нулю при больших n. Следовательно, выборочная медиана состоятельна.

17.12 Мы должны, таким образом, продолжать поиски критериев, позволяющих осуществлять выбор среди различных состоятельных оценок. Один такой критерий возникает естественным образом при рассмотрении дисперсий оценок. Дело в том, что чем меньше дисперсия оценки, тем большая доля ее распределения сосредоточена вблизи θ. Во всяком случае, это верно для распределений типа нормального. Несмещенная состоятельная оценка с меньшей дисперсией будет, следовательно, в среднем меньше отклоняться от истинного значения, чем оценка с большей дисперсией. Поэтому разумно считать ее более предпочтительной.

*) Имеется в виду значение функции плотности, отвечающее медиане. (*Прим. ред.*)

Для среднего и медианы нормальных выборок имеем из (17.3)

$$\mathsf{D}\,(\text{средн.}) = \sigma^2/n \tag{17.14}$$

(при любых n) и из **17.11**

$$\mathsf{D}\,(\text{мед.}) = \pi\sigma^2/(2n) \tag{17.15}$$

(при больших n). Поскольку $\pi/2 = 1{,}57 > 1$, то среднее является более эффективной оценкой, чем медиана, по крайней мере для больших n. Следующие данные могут быть получены с помощью таблицы XXIII из сборника *Tables for Statisticians and Biometricians, Part II*:

n	2	3	4	5
$\frac{\mathsf{D}\,(\text{мед.})}{\mathsf{D}\,(\text{сред.})}$	1,00	1,35	1,19	1,44

Таким образом, мы видим, что дисперсия выборочного среднего всегда меньше дисперсии выборочной медианы при оценивании среднего значения нормального распределения.

Пример 17.5

Мы уже знаем (см. **17.6**), что выборочное среднее не является состоятельной оценкой медианы θ распределения Коши

$$dF(x) = \frac{1}{\pi}\,\frac{dx}{\{1 + (x-\theta)^2\}}, \qquad -\infty \leqslant x \leqslant \infty.$$

В то же время, поскольку ордината медианы равна $1/\pi$, (17.13) дает для дисперсии t в случае больших выборок

$$\mathsf{D}t = \frac{\pi^2}{4n},$$

и медиана, очевидно, является состоятельной оценкой. Хотя в данном случае прямое сравнение со средним невозможно, так как дисперсия последнего не существует, тем не менее ясно, что медиана предпочтительней выборочного среднего в качестве оценки для θ. Таким образом, ситуация здесь противоположна той, которая была в случае нормального распределения. Этот факт особенно интересен, если принять во внимание сходство данных распределений.

Питмэн (1937c) дает следующее определение. Оценка t называется «более близкой» к θ, чем оценка u, если $\mathsf{P}\{|t-\theta| < |u-\theta|\} > 1/2$. Однако, как он подчеркивает, этот привлекательный критерий не обладает свойством транзитивности. См. также Гири (1944) и Джонсон (1950).

Оценки с минимальной дисперсией

17.13 Итак, кажется естественным использовать дисперсию оценки как критерий ее приемлемости. И она действительно использовалась для этой цели еще со времен Лапласа и Гаусса. Однако только относительно недавно было установлено, что при достаточно общих условиях существует нижняя граница дисперсий оценок. Для нахождения этой границы нам понадобятся некоторые предварительные результаты, которые будут полезны в дальнейшем также для других целей.

17.14 Обозначая $f(x|\theta)$ плотность непрерывной или дискретной совокупности, мы определим функцию правдоподобия *) L выборки из n независимых наблюдений равенством

$$L(x_1, x_2, \ldots, x_n|\theta) = f(x_1|\theta) f(x_2|\theta) \ldots f(x_n|\theta). \quad (17.16)$$

Поскольку L есть совместная плотность наблюдений, то

$$\int \ldots \int L\, dx_1 \ldots dx_n = 1. \quad (17.17)$$

Предположим теперь, что для всех θ существуют первые две производные по θ функции L, и продифференцируем обе части равенства (17.17). Если можно изменить порядок дифференцирования и интегрирования, например, если пределы интегрирования (т. е. область изменения x) не зависят от θ, то мы получим

$$\int \ldots \int \frac{\partial L}{\partial \theta}\, dx_1 \ldots dx_n = 0,$$

что можно переписать в виде

$$\mathsf{M}\left(\frac{\partial \log L}{\partial \theta}\right) = \int \ldots \int \left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right) L\, dx_1 \ldots dx_n = 0. \quad (17.18)$$

Дифференцирование (17.18) и последующая перемена порядка дифференцирования и интегрирования приводят к соотношению

$$\int \ldots \int \left\{\left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right)\frac{\partial L}{\partial \theta} + L\frac{\partial}{\partial \theta}\left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right)\right\} dx_1 \ldots dx_n = 0,$$

которое можно записать как

$$\int \ldots \int \left\{\left(\frac{1}{L}\frac{\partial L}{\partial \theta}\right)^2 + \frac{\partial^2 \log L}{\partial \theta^2}\right\} L\, dx_1 \ldots dx_n = 0$$

*) Р. А. Фишер называет функцию L правдоподобием, когда она рассматривается как функция от θ, и вероятностью, когда она рассматривается как функция от x при фиксированном θ. Соглашаясь с важностью проведения такого различия, мы, однако, во избежание введения дополнительных обозначений, используем в обоих случаях термин «правдоподобие» и символ L.

или же

$$\mathbf{M}\left[\left(\frac{\partial \log L}{\partial \theta}\right)^2\right] = -\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right). \tag{17.19}$$

17.15 Далее мы будем рассматривать несмещенные оценки t некоторой функции $\tau(\theta)$ от параметра θ. Это позволит рассмотреть не только смещенные и несмещенные оценки самого параметра θ, но и, например, оценивание стандартного отклонения в случае, когда параметром служит дисперсия. Итак,

$$\mathbf{M}(t) = \int \ldots \int tL\, dx_1 \ldots dx_n = \tau(\theta). \tag{17.20}$$

Дифференцирование соотношения (17.20) дает

$$\int \ldots \int t \frac{\partial \log L}{\partial \theta} L\, dx_1 \ldots dx_n = \tau'(\theta),$$

что можно переписать, учитывая (17.18), в виде

$$\tau'(\theta) = \int \ldots \int \{t - \tau(\theta)\} \frac{\partial \log L}{\partial \theta} L\, dx_1 \ldots dx_n. \tag{17.21}$$

Используя неравенство Коши — Буняковского *), из последнего соотношения получаем

$$\tau'(\theta)\}^2 \leqslant$$

$$\leqslant \int \ldots \int \{t - \tau(\theta)\}^2 L\, dx_1 \ldots dx_n \int \ldots \int \left(\frac{\partial \log L}{\partial \theta}\right)^2 L\, dx_1 \ldots dx_n$$

или, после несложных преобразований,

$$\mathbf{D}t = \mathbf{M}\{t - \tau(\theta)\}^2 \geqslant \{\tau'(\theta)\}^2 / \mathbf{M}\left[\left(\frac{\partial \log L}{\partial \theta}\right)^2\right]. \tag{17.22}$$

Это основное неравенство для дисперсии оценки. Его часто называют неравенством Крамера — Рао (Рао (1945); Крамер (1946)). Впервые оно было, по-видимому, получено Эйткином и Силверстоуном (1942). Используя (17.19), это неравенство можно переписать в следующем виде, часто более удобном на практике:

$$\mathbf{D}t \geqslant -\{\tau'(\theta)\}^2 / \mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right). \tag{17.23}$$

Мы будем называть (17.22) и (17.23) минимальной границей дисперсии (сокращенно МГД) оценки функции $\tau(\theta)$. Оценка, дисперсия которой при всех θ равна этой минимальной границе, будет называться МГД-оценкой.

*) В оригинале это неравенство называется неравенством Коши — Шварца. (*Прим. ред.*)

Для того чтобы из (17.20) следовало (17.22), необходимо лишь выполнение равенства (17.18). Если, кроме того, выполнено (17.19), то МГД может быть записана также в виде (17.23). (См. в связи с этим упражнения 17.21 и 17.22.)

17.16 Если оценивается сам параметр θ, то, заменяя в (17.22) производную $\tau'(\theta)$ ее значением $\tau'(\theta) = 1$, мы получим для *несмещенной* оценки параметра θ неравенство

$$\mathbf{D}t \geqslant 1/\mathbf{M}\left[\left(\frac{\partial \log L}{\partial \theta}\right)^2\right] = -1/\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right). \qquad (17.24)$$

Величина

$$I = \mathbf{M}\left[\left(\frac{\partial \log L}{\partial \theta}\right)^2\right] \qquad (17.25)$$

называется иногда *количеством информации*, содержащейся в выборке.

17.17 Условие, при котором достигается МГД, получить нетрудно. В самом деле, (17.22) следует непосредственно из неравенства Коши — Буняковского, а необходимым и достаточным условием того, чтобы последнее превратилось в равенство, является (см. **2.7**) пропорциональность $\{t - \tau(\theta)\}$ и $\frac{\partial \log L}{\partial \theta}$ для любых наборов наблюдений, т. е.

$$\frac{\partial \log L}{\partial \theta} = A\{t - \tau(\theta)\}. \qquad (17.26)$$

Значение A в (17.26) не зависит от наблюдений, но может быть функцией от θ, поэтому следует переписать (17.26) в виде

$$\frac{\partial \log L}{\partial \theta} = A(\theta)\{t - \tau(\theta)\}. \qquad (17.27)$$

Далее, из (17.27) и (17.18) получаем

$$\mathbf{D}\left(\frac{\partial \log L}{\partial \theta}\right) = \mathbf{M}\left[\left(\frac{\partial \log L}{\partial \theta}\right)^2\right] = \{A(\theta)\}^2 \mathbf{D}t. \qquad (17.28)$$

Так как (17.22) превращается в равенство, то подстановка в него (17.28) дает

$$\mathbf{D}t = |\tau'(\theta)/A(\theta)|. \qquad (17.29)$$

Мы заключаем отсюда, что если (17.27) выполнено, то t является МГД-оценкой для $\tau(\theta)$, причем дисперсия этой оценки дается выражением (17.29), которое равно правой части (17.23). Если $\tau(\theta) \equiv \theta$, то дисперсия t равна величине $1/A(\theta)$, которая в свою очередь, равна правой части (17.24).

Пример 17.6

Требуется оценить параметр θ в нормальной совокупности

$$dF(x) = \frac{1}{\sigma (2\pi)^{1/2}} \exp\left\{ -\frac{1}{2}\left(\frac{x-\theta}{\sigma}\right)^2 \right\} dx, \quad -\infty \leqslant x \leqslant \infty,$$

с известным σ.

Поскольку равенство

$$\frac{\partial \log L}{\partial \theta} = \frac{n}{\sigma^2}(\bar{x} - \theta)$$

имеет вид (17.27), если положить $t = \bar{x}$, $A(\theta) = n/\sigma^2$ и $\tau(\theta) = \theta$, то $\bar{x}$ будет МГД-оценкой для θ. Дисперсия этой оценки равна σ^2/n.

Пример 17.7

В случае распределения

$$dF(x) = \frac{1}{\pi} \frac{dx}{\{1 + (x-\theta)^2\}}, \quad -\infty \leqslant x \leqslant \infty,$$

для параметра θ не существует МГД-оценки, так как равенство

$$\frac{\partial \log L}{\partial \theta} = 2 \sum \frac{x-\theta}{\{1 + (x-\theta)^2\}}$$

не может быть приведено к виду (17.27).

Пример 17.8

Для распределения Пуассона

$$f(x \mid \theta) = e^{-\theta}\theta^x / x!, \quad x = 0, 1, 2, \ldots, \infty,$$

получаем

$$\frac{\partial \log L}{\partial \theta} = \frac{n}{\theta}(\bar{x} - \theta),$$

откуда следует, что $\bar{x}$ есть МГД-оценка для θ. Дисперсия этой оценки равна θ/n.

Пример 17.9

В случае биномиального распределения, для которого

$$L(r \mid \theta) = \binom{n}{r} \theta^r (1-\theta)^{n-r}, \quad r = 0, 1, 2, \ldots, n,$$

из равенства

$$\frac{\partial \log L}{\partial \theta} = \frac{n}{\theta(1-\theta)}\left(\frac{r}{n} - \theta\right)$$

следует, что r/n есть МГД-оценка параметра θ. Эта оценка имеет дисперсию, равную $\theta(1-\theta)/n$.

17.18 Из изложенного выше следует, что МГД-оценка может существовать только для какой-то одной функции $\tau(\theta)$ от параметра θ. Следующий пример иллюстрирует этот факт.

Пример 17.10

Оценить параметр θ нормального распределения

$$dF(x) = \frac{1}{\theta(2\pi)^{1/2}} \exp\left(-\frac{x^2}{2\theta^2}\right) dx, \qquad -\infty \leqslant x \leqslant \infty$$

Из равенства

$$\frac{\partial \log L}{\partial \theta} = -\frac{n}{\theta} + \frac{\sum x^2}{\theta^2} = \frac{n}{\theta^3}\left(\frac{1}{n}\sum x^2 - \theta^2\right)$$

немедленно следует, что $\frac{1}{n}\sum x^2$ есть МГД-оценка для генеральной дисперсии θ^2, причем дисперсия этой оценки, согласно (17.29), равна $\frac{\theta^3}{n} \cdot \frac{d}{d\theta}(\theta^2) = \frac{2\theta^4}{n}$. Однако для самого θ МГД-оценки в данном случае не существует.

Уравнение (17.27) дает условие, которому должна удовлетворять плотность распределения, для того чтобы для некоторой функции $\tau(\theta)$ от параметра θ существовала МГД-оценка. Но если это условие не выполняется, то еще может существовать оценка, которая будет равномерно по θ иметь дисперсию, меньшую дисперсии любой другой оценки. Такую оценку мы назовем оценкой с минимальной дисперсией (МД-оценка). Другими словами, минимальная *достижимая* дисперсия может быть больше нижней границы дисперсии, даваемой неравенством Крамера — Рао. Когда же условия регулярности, приводящие к неравенству Крамера — Рао, не выполнены, минимальная достижимая дисперсия может быть и меньше нижней границы, получающейся при формальном применении этого неравенства. Однако в любом случае из равенства (17.27) вытекает, что может существовать только одна функция от θ, для которой достигается МГД, а именно функция, являющаяся математическим ожиданием статистики t, через которую линейно выражается производная логарифма функции правдоподобия $\partial \log L/\partial\theta$.

17.19 Интегрируя соотношение (17.27), определяющее вид функции правдоподобия, и обозначая интеграл от произвольной функции $A(\theta)$ снова через $A(\theta)$, мы получим

$$\log L = tA(\theta) + P(\theta) + R(x_1, x_2, \ldots, x_n).$$

Отсюда следует, что плотность распределения должна иметь вид

$$f(x\,|\,\theta) = \exp\{A(\theta)\,B(x) + C(x) + D(\theta)\}, \qquad (17.30)$$

где

$$t=\sum_{i=1}^{n} B(x_i), \qquad R(x_1, \ldots, x_n)=\sum_{i=1}^{n} C(x_i) \quad \text{и} \quad P(\theta)=nD(\theta).$$

Семейство распределений вида (17.30) часто называют экспоненциальным. Мы возвратимся к нему в **17.36**.

17.20 Когда МГД (17.22) не достигается, можно найти более точную нижнюю границу дисперсии оценки, т. е. границу, бóльшую той, которая дается неравенством Крамера — Рао. Существенным условием достижимости (17.22) является линейная зависимость разности $t-\tau(\theta)$ от $\frac{\partial \log L}{\partial \theta}=\frac{1}{L}\frac{\partial L}{\partial \theta}$. Однако если такой оценки t, для которой это условие выполняется, не существует, может еще существовать оценка, для которой $t-\tau(\theta)$ будет линейной функцией от $\frac{1}{L}\frac{\partial L}{\partial \theta}$, $\frac{1}{L}\frac{\partial^2 L}{\partial \theta^2}$ и от более высоких производных функции правдоподобия. Этот факт приводит к следующему результату, принадлежащему Бхаттачария (1946).

Пусть функция $\tau(\theta)$ оценивается статистикой t. Введем обозначения

$$L^{(r)}=\frac{\partial^r L}{\partial \theta^r} \quad \text{и} \quad \tau^{(r)}=\frac{\partial^r \tau(\theta)}{\partial \theta^r}$$

и определим функцию

$$D_s=t-\tau(\theta)-\sum_{r=1}^{s} a_r L^{(r)}/L, \tag{17.31}$$

где a_r — константы, которые будут определены ниже. Дифференцирование (17.17) r раз дает (при допустимости перемены порядка дифференцирования и интегрирования)

$$\mathsf{M}\left(L^{(r)}/L\right)=0, \tag{17.32}$$

откуда

$$\mathsf{M}(D_s)=0. \tag{17.33}$$

Дисперсия D_s равна, следовательно,

$$\mathsf{M}\left(D_s^2\right)=\int \ldots \int \left\{(t-\tau(\theta))-\sum_r a_r L^{(r)}/L\right\}^2 L\, dx_1 \ldots dx_n. \tag{17.34}$$

Минимизация (17.34) по a_r приводит к системе

$$\int \ldots \int \left\{(t-\tau(\theta))-\sum_r a_r L^{(r)}/L\right\} \frac{L^{(p)}}{L} \cdot L\, dx_1 \ldots dx_n=0 \tag{17.35}$$

($p = 1, 2, \ldots, s$), которую можно переписать следующим образом:

$$\int \ldots \int (t - \tau(\theta)) L^{(p)} dx_1 \ldots dx_n =$$

$$= \sum_r a_r \int \ldots \int \frac{L^{(r)}}{L} \cdot \frac{L^{(p)}}{L} L \, dx_1 \ldots dx_n. \quad (17.36)$$

Левая часть (17.36), согласно (17.32) и (17.20), равна

$$\int \ldots \int t L^{(p)} dx_1 \ldots dx_n = \tau^{(p)}.$$

Правую же можно записать в виде

$$\sum a_r \mathsf{M}\left(\frac{L^{(r)}}{L} \cdot \frac{L^{(p)}}{L}\right).$$

Подставляя эти значения в (17.36), получим

$$\tau^{(p)} = \sum_{r=1}^{s} a_r \mathsf{M}\left(\frac{L^{(r)}}{L} \cdot \frac{L^{(p)}}{L}\right), \quad p = 1, 2, \ldots, s. \quad (17.37)$$

Если матрица $J_{rp} = \mathsf{M}\left(\frac{L^{(r)}}{L} \cdot \frac{L^{(p)}}{L}\right)$ этой линейной системы невырождена, то система имеет решение

$$a_r = \sum_{p=1}^{s} \tau^{(p)} J_{rp}^{-1}, \quad r = 1, 2, \ldots, s. \quad (17.38)$$

Следовательно, если (17.34) минимально, то (17.31) принимает значение

$$D_s = t - \tau(\theta) - \sum_{r=1}^{s} \sum_{p=1}^{s} \tau^{(p)} J_{rp}^{-1} L^{(r)} / L. \quad (17.39)$$

Само же (17.34), как следует из (17.39), равно

$$\mathsf{M}(D_s^2) = \int \ldots \int \left\{ (t - \tau(\theta)) - \sum_r \sum_p \tau^{(p)} J_{rp}^{-1} L^{(r)} / L \right\}^2 L \, dx_1 \ldots dx_n.$$

Это выражение, учитывая (17.35), можно записать в виде

$$\mathsf{M}(D_s^2) = \int \ldots \int \{t - \tau(\theta)\}^2 L \, dx_1 \ldots dx_n -$$

$$- \sum_r \sum_p \tau^{(p)} J_{rp}^{-1} \int \ldots \int t L^{(r)} dx_1 \ldots dx_n$$

или, окончательно,

$$\mathsf{M}(D_s^2) = \mathsf{D}t - \sum_r \sum_p \tau^{(p)} J_{rp}^{-1} \tau^{(r)}. \quad (17.40)$$

Так как левая часть (17.40) неотрицательна, то мы получаем нужное нам неравенство

$$\mathsf{D}t \geqslant \sum_{r=1}^{s} \sum_{p=1}^{s} \tau^{(p)} J_{rp}^{-1} \tau^{(r)}. \tag{17.41}$$

В частном случае $s = 1$ неравенство (17.41) превращается в неравенство (17.22).

17.21 Условием достижения нижней границы в (17.41) является условие $\mathsf{M}(D_s^2) = 0$, которое ввиду (17.33) сводится к $D_s = 0$, что дает (см. (17.39))

$$t - \tau(\theta) = \sum_{r=1}^{s} \sum_{p=1}^{s} \tau^{(p)} J_{rp}^{-1} L^{(r)} / L. \tag{17.42}$$

Условие (17.42), требующее линейной зависимости $t - \tau(\theta)$ от величин $L^{(r)}/L$, является обобщением (17.27). Если $t - \tau(\theta)$ есть линейная функция от первых s таких величин, то, конечно, мы ничего не выиграем, добавляя члены более высокого порядка. С другой стороны, правая часть (17.41) является неубывающей функцией от s, поскольку минимальная дисперсия функции D_s не может возрастать при добавлении новых коэффициентов a_r. Таким образом, если при оценивании заданной функции $\tau(\theta)$ не существует МГД-оценки, то мы можем надеяться найти оценку, дисперсия которой будет равна нижней границе, даваемой (17.41) при некотором $s > 1$.

17.22 Посмотрим теперь, как изменится нижняя граница дисперсии, если в (17.41) взять $s = 2$ вместо $s = 1$. Используя обозначение

$$J_{rp} = \mathsf{M}\left(\frac{L^{(r)}}{L} \cdot \frac{L^{(p)}}{L}\right), \tag{17.43}$$

можно записать (17.41) в виде

$$\begin{vmatrix} \mathsf{D}t & \tau' & \tau'' \\ \tau' & J_{11} & J_{12} \\ \tau'' & J_{12} & J_{22} \end{vmatrix} \geqslant 0 \tag{17.44}$$

или, что то же самое,

$$\mathsf{D}t \geqslant \frac{(\tau')^2}{J_{11}} + \frac{(\tau' J_{12} - \tau'' J_{11})^2}{J_{11}(J_{11}J_{22} - J_{12}^2)}. \tag{17.45}$$

Второе слагаемое в правой части (17.45) характеризует полученное уточнение границы. Записывая $J_{rp}(n)$ в виде функции от объема выборки и пользуясь соотношениями

$$\left.\begin{aligned} J_{11}(n) &= nJ_{11}(1), \\ J_{12}(n) &= nJ_{12}(1), \\ J_{22}(n) &= nJ_{22}(1) + 2n(n-1)\{J_{11}(1)\}^2, \end{aligned}\right\} \tag{17.46}$$

получим

$$\mathrm{D}t \geqslant \frac{(\tau')^2}{nJ_{11}(1)} + \frac{\{\tau'' - \tau' J_{12}(1)/J_{11}(1)\}^2}{2n^2\{J_{11}(1)\}^2} + o\left(\frac{1}{n^2}\right), \qquad (17.47)$$

откуда следует, что поправка к границе имеет порядок $1/n^2$, в то время как главный член имеет порядок $1/n$. Следовательно, величина поправки важна только при $\tau' = 0$, т. е. когда главный член исчезает. В случае, когда t является оценкой для самого θ, (17.47) принимает вид

$$\mathrm{D}t \geqslant \frac{1}{nJ_{11}(1)} + \frac{\{J_{12}(1)\}^2}{2n^2\{J_{11}(1)\}^4} + o\left(\frac{1}{n^2}\right),$$

что можно переписать, используя (17.46) и возвращаясь к первоначальным обозначениям, следующим образом:

$$\mathrm{D}t \geqslant \frac{1}{J_{11}} + \frac{J_{12}^2}{2J_{11}^4} + o\left(\frac{1}{n^2}\right). \qquad (17.48)$$

Если граница при $s = 2$ равна границе при $s = 1$, то из этого, вообще, не следует, что последняя достижима. Однако для экспоненциального семейства это так (см. Патил и Шоррок (1965)).

Пример 17.11

Мы уже знаем, что несмещенной оценкой параметра θ биномиального распределения является статистика $r/n = p$, а несмещенной оценкой для θ^2 — статистика $r(r-1)/\{n(n-1)\}$, поэтому естественной несмещенной оценкой для функции $\theta(1-\theta)$ будет

$$t = \frac{r}{n} - \frac{r(n-1)}{n(n-1)} = \left(\frac{n}{n-1}\right) p(1-p).$$

Дисперсия t в случае больших выборок с точностью до членов первого порядка дается (см. (10.14)) выражением

$$\mathrm{D}t \sim \left(\frac{n}{n-1}\right)^2 \left\{\frac{d}{dp}[p(1-p)]\right\}^2_{p=\theta} \mathrm{D}p \sim \frac{\theta(1-\theta)(1-2\theta)^2}{n},$$

которое совпадает с первым членом в (17.41) и (17.45), поскольку $\tau' = (1-2\theta)$ и $J_{11} = n/\{\theta(1-\theta)\}$. Таким образом, если этот член не равен нулю, то t будет МГД-оценкой (асимптотически). Однако при $\theta = 1/2$ он равен нулю, поэтому мы должны рассмотреть следующий член, чтобы получить ненулевую границу дисперсии. Легко проверить, что

$$J_{12} = -\frac{2n(1-2\theta)}{\theta^2(1-\theta)^2},$$

следовательно, второй член в (17.47) равен

$$\frac{2}{n^2}\{(1-2\theta)^2 - \theta(1-\theta)\}^2.$$

При $\theta = 1/2$ этот член равен $1/(8n^2)$. Записав статистику t в виде

$$t = \left(\frac{n}{n-1}\right)\left\{\frac{1}{4} - \left(p - \frac{1}{2}\right)^2\right\},$$

можно найти точную дисперсию t при $\theta = 1/2$, используя моменты биномиального распределения. Эта дисперсия равна $\mathsf{D}t = 1/\{8n(n-1)\}$, т. е. совпадает со вторым слагаемым в (17.47) до членов порядка $1/n^2$.

17.23 Пример 17.11 выявляет один довольно важный факт. Если для функции $\tau(\theta)$ существует МГД-оценка t, т. е.

$$\mathsf{D}t = \{\tau'(\theta)\}^2/I_{11}, \tag{17.49}$$

и требуется оценить некоторую другую функцию, скажем $\psi\{\tau(\theta)\}$, то из известных асимптотических результатов и из (17.49) следует (при условии, что ψ' не равна нулю), что

$$\mathsf{D}\{\psi(t)\} \sim \left(\frac{\partial\psi}{\partial\tau}\right)^2 \mathsf{D}t \sim \left(\frac{\partial\psi}{\partial\tau}\right)^2 \left(\frac{\partial\tau}{\partial\theta}\right)^2 \Big/ I_{11},$$

или

$$\mathsf{D}\{\psi(t)\} \sim \left(\frac{\partial\psi}{\partial\theta}\right)^2 \Big/ I_{11}.$$

Другими словами, из того, что для некоторой функции от θ существует МГД-оценка, следует, что для любой функции от θ существует оценка с дисперсией, асимптотически равной МГД. Более того, этой оценкой служит соответствующая функция от t. В **17.35** мы получим, в дополнение к сказанному, один более точный результат, касающийся функций от МГД-оценок.

17.24 Не представляет никакого труда распространить результаты этой главы относительно МГД-оценок на случай нескольких параметров $\theta_1, \theta_2, \ldots, \theta_k$. Если ставится задача об оценивании некоторой функции от этих параметров, скажем $\tau(\theta_1, \theta_2, \ldots, \theta_k)$, то аналогом самого простого результата, (17.22), будет неравенство

$$\mathsf{D}t \geqslant \sum_{i=1}^{k}\sum_{j=1}^{k} \frac{\partial\tau}{\partial\theta_i}\frac{\partial\tau}{\partial\theta_j} I_{ij}^{-1}, \tag{17.50}$$

где I_{ij}^{-1} — элементы матрицы, обратной к

$$\{I_{ij}\} = \left\{\mathsf{M}\left\{\frac{1}{L}\frac{\partial L}{\partial\theta_i}\cdot\frac{1}{L}\frac{\partial L}{\partial\theta_j}\right\}\right\}.$$

Как и в одномерном случае, (17.50) учитывает только члены порядка $1/n$; неравенство же, учитывающее члены более высокого порядка, имеет более сложный вид.

17.25 Оказывается, можно получить даже более точные, чем рассмотренные выше, нижние границы для дисперсий оценок, причем без условий регулярности. Следуя Киферу (1952), рассмотрим случай несмещенных оценок.

Обозначим, как и в (17.16), через $L(x|\theta)$ функцию правдоподобия, а через h — случайную величину, независимую от x и определенную так, что сумма $(\theta + h)$ не выходит за пределы допустимых значений для θ. Пусть $\lambda_1(h)$ и $\lambda_2(h)$ — произвольные функции распределения. Тогда, если t — несмещенная оценка для θ, то можно написать

$$\int \{t - (\theta + h)\} L(x|\theta + h)\, dx = 0,$$

используя один знак интеграла и один дифференциал dx для обозначения n-кратного интегрирования. Интегрируя еще раз по $\lambda_i(h)$ $(i = 1, 2)$, получим

$$\int \left[\int \{t - (\theta - h)\} L(x|\theta + h)\, dx \right] d\lambda_i(h) = 0.$$

Таким образом, можно написать

$$\mathbf{M}_1(h) - \mathbf{M}_2(h) = \int (t - \theta) \left[\int L(x|\theta + h)\, d\{\lambda_1(h) - \lambda_2(h)\} \right] dx =$$

$$= \int (t - \theta) \{L(x|\theta)\}^{1/2} \left\{ \frac{\int L(x|\theta + h)\, d(\lambda_1 - \lambda_2)}{\{L(x|\theta)\}^{1/2}} \right\} dx.$$

Пользуясь теперь неравенством Коши — Буняковского, получим

$$\mathbf{M}_1(h) - \mathbf{M}_2(h))^2 \leqslant \int (t - \theta)^2 L(x|\theta)\, dx \int \left\{ \frac{\int L(x|\theta + h)\, d(\lambda_1 - \lambda_2)}{\{L(x|\theta)\}^{1/2}} \right\}^2 dx,$$

что дает

$$\mathbf{D}t \geqslant \frac{\{\mathbf{M}_1(h) - \mathbf{M}_2(h)\}^2}{\int \left[\left\{ \int L(x|\theta + h)\, d(\lambda_1 - \lambda_2) \right\}^2 \Big/ L(x|\theta) \right] dx}.$$

Поскольку это верно для любых λ_1 и λ_2, то мы имеем

$$\mathbf{D}t \geqslant \sup_{\lambda_1, \lambda_2} \frac{\{\mathbf{M}_1(h) - \mathbf{M}_2(h)\}^2}{\int \left[\left\{ \int L(x|\theta + h)\, d(\lambda_1 - \lambda_2) \right\}^2 \Big/ L(x|\theta) \right] dx}. \qquad (17.51)$$

Верхняя грань в (17.51) берется по всем $\lambda_1 \neq \lambda_2$, для которых определено подинтегральное выражение в знаменателе. Баранкин (1949) доказал, что неравенства этого типа дают наилучшие возможные границы для $\mathbf{D}\,t$.

Если распределение $\lambda_2(h)$ сосредоточено в точке $h \neq 0$, то (17.51) принимает вид

$$\mathbf{D}t \geqslant \sup_{\lambda_1} \frac{\{\mathbf{M}_1(h)\}^2}{\int \left[\left\{ \int L(x|\theta + h)\, d\lambda_1 \right\}^2 \Big/ L(x|\theta) \right] dx - 1}. \qquad (17.52)$$

Если, кроме того, распределение λ_1 также сосредоточено в некоторой точке $h \neq 0$, то неравенство (17.52) превращается в

$$\mathbf{D}t \geqslant \frac{1}{\inf\limits_h (1/h^2) \left\{ \int [(L(x|\theta + h))^2 / L(x|\theta)]\, dx - 1 \right\}}. \qquad (17.53)$$

Нижняя грань берется по всем $h \neq 0$, для которых из $L(x|\theta) = 0$ следует $L(x|\theta + h) = 0$. Неравенство (17.53), установленное Чепмэном и Роббинсом (1951), дает, по меньшей мере, такую же хорошую границу, как и (17.24), но применимо в более общем случае. Знаменатель в правой части (17.24), равный

$$\mathsf{M}\left[\left(\frac{\partial \log L}{\partial \theta}\right)^2\right] = \int \lim_{h\to 0}\left[\frac{L(x|\theta+h) - L(x|\theta)}{hL(x|\theta)}\right]^2 L(x|\theta)\,dx,$$

в случае допустимости перемены порядка интегрирования и перехода к пределу можно записать в виде

$$\mathsf{M}\left[\left(\frac{\partial \log L}{\partial \theta}\right)^2\right] = \lim_{h\to 0}\frac{1}{h^2}\left[\int \frac{\{L(x|\theta+h)\}^2}{L(x|\theta)}\,dx - 1\right].$$

Знаменатель в правой части (17.53) есть нижняя грань этой величины по всем допустимым значениям h и не может, следовательно, быть больше предела этого выражения при $h \to 0$, т. е. (17.53) дает, по меньшей мере, такие же хорошие результаты, как и (17.24). Отсюда следует, что (17.52) и (17.51) дают, вообще говоря, еще лучшие результаты, однако, чтобы их использовать, нужно в каждом конкретном случае находить λ_1 и λ_2. Неравенство же (17.53) можно применять сразу. В случае, когда область определения распределения зависит от оцениваемого параметра и неравенство (17.24) неприменимо, более эффективным средством изучения свойств наилучших оценок является использование достаточных статистик, которые будут рассмотрены несколько позже.

17.26 До сих пор мы в основном занимались оценками, дисперсия которых равномерно по θ достигала МГД. Однако во многих случаях таких оценок не существует, даже когда выполнены условия, при которых была получена эта граница. Остается надежда, что тогда существует МД-оценка, т. е. оценка, имеющая для всех θ минимальную достижимую дисперсию. Единственна ли такая оценка? Легко показать, что если МД-оценка существует, то она всегда единственна, безотносительно к достижению какой-либо границы.

Пусть t_1 и t_2 — несмещенные МД-оценки для $\tau(\theta)$, и пусть дисперсия этих оценок равна V. Рассмотрим новую оценку для $\tau(\theta)$

$$t_3 = \tfrac{1}{2}(t_1 + t_2)$$

с дисперсией

$$\mathsf{D}t_3 = \tfrac{1}{4}\{\mathsf{D}t_1 + \mathsf{D}t_2 - 2\operatorname{cov}(t_1, t_2)\}. \qquad (17.54)$$

Согласно неравенству Коши — Буняковского

$$\operatorname{cov}(t_1, t_2) = \int \ldots \int (t_1 - \tau)(t_2 - \tau) L\,dx_1 \ldots dx_n \leqslant$$

$$\leqslant \left\{\int \ldots \int (t_1 - \tau)^2 L\,dx_1 \ldots dx_n \int \ldots \int (t_2 - \tau)^2 L\,dx_1 \ldots dx_n\right\}^{1/2} \leqslant$$

$$\leqslant (\mathsf{D}t_1\,\mathsf{D}t_2)^{1/2} \leqslant V, \qquad (17.55)$$

откуда

$$\mathsf{D}t_3 \leqslant V.$$

Поскольку t_1 и t_2 имеют по предположению минимальную дисперсию, то в последнем неравенстве, а следовательно, и в (17.55) должен стоять знак равенства. Для этого $t_1 - \tau$ и $t_2 - \tau$ должны быть пропорциональны:

$$(t_1 - \tau) = k(\theta)(t_2 - \tau). \tag{17.56}$$

Но если это так, то

$$\operatorname{cov}(t_1, t_2) = k(\theta)\,\mathsf{D}t_2 = k(\theta)\,V,$$

откуда ввиду равенства в (17.55)

$$k(\theta) = 1.$$

Таким образом, t_1 совпадает с t_2, т. е. МД-оценка единственна.

17.27 Обобщая доказательство предыдущего пункта, можно получить одно интересное неравенство для коэффициента корреляции ρ между двумя оценками, который согласно **16.23** определяется как отношение их ковариации к квадратному корню из произведения дисперсий. Из неравенства (17.55) следует, что всегда $\rho^2 \leqslant 1$.

Пусть t — несмещенная оценка для $\tau(\theta)$ с минимальной дисперсией V, а t_1 и t_2 — любые другие несмещенные оценки для $\tau(\theta)$. Рассмотрим новую оценку для $\tau(\theta)$

$$t_3 = at_1 + (1 - a)t_2,$$

которая имеет дисперсию

$$\mathsf{D}t_3 = a^2\,\mathsf{D}t_1 + (1 - a)^2\,\mathsf{D}t_2 + 2a(1 - a)\operatorname{cov}(t_1, t_2). \tag{17.57}$$

Вводя обозначения

$$\mathsf{D}t_1 = k_1 V, \qquad \mathsf{D}t_2 = k_2 V, \qquad k_1,\ k_2 > 1,$$

можно записать (17.57) в виде

$$\mathsf{D}t_3/V = a^2 k_1 + (1 - a)^2 k_2 + 2a(1 - a)\operatorname{cov}(t_1, t_2)/V \tag{17.58}$$

или, обозначая

$$\rho = \frac{\operatorname{cov}(t_1, t_2)}{(\mathsf{D}t_1 \mathsf{D}t_2)^{1/2}} = \frac{\operatorname{cov}(t_1, t_2)}{(k_1 k_2)^{1/2} V},$$

в виде

$$\mathsf{D}t_3/V = a^2 k_1 + (1 - a)^2 k_2 + 2a(1 - a)\rho\,(k_1 k_2)^{1/2},$$

что можно переписать, учитывая неравенство $\mathsf{D}t_3 \geqslant V$, следующим образом:

$$a^2 k_1 + (1 - a)^2 k_2 + 2a(1 - a)\rho\,(k_1 k_2)^{1/2} \geqslant 1. \tag{17.59}$$

Неравенство (17.59) является квадратичным неравенством по a:

$$a^2\{k_1+k_2-2\rho(k_1k_2)^{1/2}\}+2a\{\rho(k_1k_2)^{1/2}-k_2\}+(k_2-1)\geqslant 0.$$

Поскольку корни соответствующего уравнения могут быть только мнимыми или равными, то его дискриминант неположителен:

$$\{\rho(k_1k_2)^{1/2}-k_2\}^2\leqslant\{k_1+k_2-2\rho(k_1k_2)^{1/2}\}(k_2-1),$$

или

$$\{\rho(k_1k_2)^{1/2}-1\}^2\leqslant(k_1-1)(k_2-1).$$

Следовательно,

$$\frac{1}{(k_1k_2)^{1/2}}+\left\{\frac{(k_1-1)(k_2-1)}{k_1k_2}\right\}^{1/2}\geqslant\rho\geqslant\frac{1}{(k_1k_2)^{1/2}}-\left\{\frac{(k_1-1)(k_2-1)}{k_1k_2}\right\}^{1/2}.$$

Обозначая $E_1=1/k_1$, $E_2=1/k_2$, имеем окончательно

$$(E_1E_2)^{1/2}-\{(1-E_1)(1-E_2)\}^{1/2}\leqslant\rho\leqslant(E_1E_2)^{1/2}+\{(1-E_1)(1-E_2)\}^{1/2}. \tag{17.60}$$

Когда E_1 или $E_2=1$, т. е. t_1 или t_2 есть МД-оценка для $\tau(\theta)$, неравенство (17.60) превращается в равенство

$$\rho=E^{1/2}, \tag{17.61}$$

где E — обратная величина дисперсии другой оценки, называемая обычно по причинам, которые будут рассмотрены в **17.28—29**, ее эффективностью.

Эффективность

17.28 До сих пор мы занимались точными результатами, относящимися к МД-оценкам, в том смысле, что на объем выборки не накладывалось никаких ограничений. Теперь мы переходим к рассмотрению свойств, справедливых лишь для больших выборок. Дело в том, что если даже МД-оценка не существует ни для какого значения n, то такая оценка может существовать для $n\to\infty$. Большинство оценок, с которыми мы имеем дело, в силу центральной предельной теоремы распределены асимптотически нормально; поэтому при больших выборках распределения таких оценок будут характеризоваться только двумя параметрами — средним и дисперсией. Если оценка состоятельна, то она, как правило, будет асимптотически несмещенной (см. **17.9**). Поэтому основным средством сравнения состоятельных, асимптотически нормальных оценок некоторой функции от параметра становится дисперсия.

Оценка такого рода, имеющая минимальную дисперсию для больших выборок, называется *эффективной* *). Этот термин был введен Фишером (1921a). Из результатов **17.26** следует, что все эффективные оценки асимптотически эквивалентны.

17.29 В случае асимптотически нормальных оценок и больших n, в отличие от случая малых n и оценок с произвольным распределением, разумно ввести *меру* эффективности оценки. Мы определим эффективность некоторой оценки по отношению к эффективной оценке как величину, обратную отношению объемов выборок, приводящих к одной и той же дисперсии оценки, т. е. дающих одинаковую точность.

Если t_1 — эффективная оценка, а t_2 — некоторая другая оценка и если (как это обычно бывает) дисперсия для большой выборки есть простая функция от величины, обратной объему выборки, то нашему определению эффективности можно придать простую форму. Пусть

$$\left.\begin{aligned} V_1 &= \mathsf{D}(t_1 \mid n_1) \sim a_1/n_1^r \qquad (r>0), \\ V_2 &= \mathsf{D}(t_2 \mid n_2) \sim a_2/n_2^s \qquad (s>0), \end{aligned}\right\} \tag{17.62}$$

где a_1, a_2 — константы, не зависящие от n, а под знаком дисперсии указан объем выборки. Чтобы получить $V_1 = V_2$, мы должны иметь

$$1 = \frac{V_1}{V_2} = \lim_{n_1,\, n_2 \to \infty} \frac{a_1}{a_2}\frac{n_2^s}{n_1^r},$$

т. е.

$$\frac{a_2}{a_1} = \lim \frac{n_2^s}{n_1^r} = \lim \left(\frac{n_2}{n_1}\right)^s \cdot \frac{1}{n_1^{r-s}}. \tag{17.63}$$

Так как t_1 по предположению эффективна, то $r \geqslant s$. При $r > s$ второй множитель в правой части (17.63) стремится к нулю; следовательно, чтобы произведение оставалось равным a_2/a_1, должно выполняться условие

$$\frac{n_2}{n_1} \to \infty, \qquad r > s, \tag{17.64}$$

которое показывает, что эффективность t_2 равна нулю. При $r = s$ имеем равенство

$$\lim \frac{n_2}{n_1} = \left(\frac{a_2}{a_1}\right)^{1/r},$$

которое в силу (17.62) может быть записано в виде

$$\lim \frac{n_2}{n_1} = \lim \left(\frac{V_2}{V_1}\right)^{1/r}.$$

*) См., однако, обсуждение «сверхэффективности» в **18.17**.

Эффективность t_2 обратна этой величине и равна

$$E = \lim \left(\frac{V_1}{V_2}\right)^{1/r}. \tag{17.65}$$

Заметим, что при $r > s$ (17.65) дает тот же результат, что и (17.64). Если $r = 1$, как это чаще всего бывает, то (17.65) сводится к величине, обратной отношению дисперсий. С этим случаем мы встречались в **17.27**. Таким образом, при сравнении оценок с дисперсиями порядка $1/n$ эффективность по отношению к эффективной оценке обратна отношению дисперсий.

Если дисперсия эффективной оценки не имеет такой простой формы, как в (17.62), то измерение относительной эффективности становится более трудной задачей (см. упражнение 18.21).

Пример 17.12

В примере 17.6 мы показали, что выборочное среднее является МГД-оценкой для среднего μ нормальной совокупности с дисперсией σ^2/n. Тем более оно будет эффективной оценкой. В примере 11.12 было показано, что выборочное среднее имеет точное нормальное распределение. В **17.11—12** мы видели, что выборочная медиана распределена асимптотически нормально со средним μ и дисперсией $\pi\sigma^2/(2n)$. Таким образом, из (17.65) с $r = 1$ следует, что эффективность выборочной медианы равна $2/\pi = 0{,}637$.

Пример 17.13

Оценка с большей эффективностью, без сомнения, предпочтительней оценки с меньшей эффективностью, если другие характеристики этих оценок одинаковы. Но эти характеристики могут быть и разными. Так, иногда получение эффективной оценки t_1 требует больше вычислений, чем получение некоторой другой оценки t_2. Трудности, связанные с увеличением объема вычислений, могут в этом случае свести на нет выигрыш от уменьшения числа наблюдений, особенно когда легко получить дополнительные наблюдения.

Рассмотрим задачу оценивания стандартного отклонения нормальной совокупности с неизвестной дисперсией σ^2 и неизвестным средним. Двумя возможными оценками стандартного отклонения являются выборочное стандартное отклонение и выборочное среднее отклонение, умноженное на $(\pi/2)^{1/2}$ (см. **5.26**). Вторая оценка, как правило, вычисляется легче, и если в нашем распоряжении имеется много наблюдений (например, ищется стандартное отклонение показаний барометра и есть

запись большого числа этих показаний), то может оказаться, что лучше пользоваться средним отклонением, а не стандартным отклонением. (Обе оценки асимптотически нормальны.)

В случае больших выборок дисперсия среднего отклонения равна $\frac{\sigma^2}{n}\left(1-\frac{2}{\pi}\right)$ (см. (10.39)). Дисперсия соответствующей оценки для σ получается умножением этой величины на $\pi/2$:

$$\frac{\pi}{2}\cdot\frac{\sigma^2}{n}\left(1-\frac{2}{\pi}\right)=\frac{\sigma^2}{2n}(\pi-2).$$

Стандартное отклонение имеет дисперсию $\sigma^2/2n$ (см. **10.8**(d)), причем дальше мы увидим, что оно является эффективной оценкой. Таким образом, можно вычислить эффективность первой оценки

$$E=\frac{\sigma^2}{2n}\Big/\left\{\frac{\sigma^2}{2n}(\pi-2)\right\}=\frac{1}{\pi-2}=0{,}876.$$

Точность оценки, основанной на среднем отклонении и вычисленной по выборке объема 1000, следовательно, примерно равна точности стандартного отклонения, вычисленного по выборке объема 876. Если легче вычислить среднее отклонение 1000 наблюдений, чем стандартное отклонение 876 наблюдений, и недостатка в наблюдениях нет, то может оказаться предпочтительней взять первую оценку.

Однако надо помнить, что, используя такую процедуру, мы сознательно соглашаемся на потерю части информации. Приложив дополнительные усилия, мы могли бы увеличить эффективность нашей оценки с 0,876 до единицы, или примерно на 14 процентов от первоначальной величины.

Оценивание по методу минимума среднего квадрата ошибки

17.30 Мы рассматривали вопросы о несмещенности и о минимизации выборочной дисперсии более или менее независимо друг от друга. Но иногда их уместно изучать одновременно. Смещение не обязательно должно играть бóльшую роль, чем дисперсия. Что мы в действительности требуем от оценки t — это чтобы она была «близка» к истинному значению θ. Рассмотрим поэтому вместо ее среднего квадрата ошибки относительно собственного математического ожидания средний квадрат ошибки относительно истинного значения. Получаем

$$\mathsf{M}(t-\theta)^2=\mathsf{M}\{(t-\mathsf{M}(t))+(\mathsf{M}(t)-\theta)\}^2=\mathsf{D}t+\{\mathsf{M}(t)-\theta\}^2$$

(член с удвоенным произведением равен нулю). Второе слагаемое в правой части есть просто квадрат смещения t при оценивании θ. Если t — несмещенная оценка, то это слагаемое равно нулю и средний квадрат ошибки совпадает с дисперсией.

Однако в общем случае минимизация среднего квадрата ошибки дает результаты, отличающиеся от тех, которые получаются при минимизации дисперсии.

Пример 17.14

На какое число нужно умножить выборочное среднее $\bar{x}$, чтобы получить для генерального среднего оценку с наименьшим средним квадратом ошибки? Из предыдущих результатов имеем $\mathsf{M}(a\bar{x}) = a\mu$, $\mathsf{D}(a\bar{x}) = a^2\sigma^2/n$, где σ^2 — генеральная дисперсия, а n — объем выборки. Следовательно,

$$\mathsf{M}(a\bar{x} - \mu)^2 = a^2\sigma^2/n + \mu^2(a - 1)^2.$$

Минимизация последнего выражения по a приводит к уравнению

$$2a\sigma^2/n + 2\mu^2(a - 1) = 0,$$

из которого получаем

$$a = \frac{\mu^2}{\mu^2 + \sigma^2/n}.$$

Если $n \to \infty$, то $a \to 1$, и мы получаем несмещенную оценку; но при любом конечном n $a > 1$.

При известной функциональной связи между μ и σ^2 можно пойти дальше. Например, если $\sigma^2 = \mu^2$, то $a = n/(n + 1)$.

Очевидно, рассмотрения такого рода представляют ценность только в случае выборки из распределения, параметры которого связаны известной функциональной зависимостью. Оценивание по методу минимума среднего квадрата ошибки применяется не очень часто, но надо заметить, что причиной этого являются скорее практические трудности, чем теоретические соображения. В некотором смысле успешное изучение несмещенных МД-оценок связано с тем, что требование несмещенности позволяет обойти эти трудности.

Достаточные статистики

17.31 Рассмотренные нами критерии, именно несмещенность, состоятельность, минимальность дисперсии и эффективность, дают довольно разумные характеристики оценок. Для их более глубокого изучения нам потребуется понятие *достаточности*, введенное Фишером (1921а, 1925).

Пусть сначала оценивается один параметр θ. Имеется бесчисленное множество возможных оценок для θ, и мы должны выбрать какую-то из них. Рассмотрим для выборки объема $n \geqslant 2$ совместное распределение r функционально независимых статистик

$$f_r(t, t_1, \ldots, t_{r-1} \mid \theta), \qquad r = 2, 3, \ldots, n,$$

из которых нас особо интересует статистика t. Используя теорему умножения вероятностей (7.9), можно записать это распределение как произведение маргинального распределения t на условное распределение других статистик при фиксированном t:

$$f_r(t, t_1, \ldots, t_{r-1} \mid \theta) = g(t \mid \theta)\, h_{r-1}(t_1, \ldots, t_{r-1} \mid t, \theta). \quad (17.66)$$

Если второй сомножитель в правой части (17.66) *не зависит* от θ, то очевидно, что имеет место ситуация, когда знание статистик $t_1, \ldots, t_{r-1}$ ничего не добавляет к знаниям о параметре θ, полученным на основании статистики t. Если, далее, это верно для каждого r и для *любого* набора $(r-1)$ статистик t_i, то можно с полным основанием сказать, что t содержит всю информацию относительно θ, имеющуюся в выборке. Статистика t будет в этом случае называться *достаточной* для θ. Таким образом статистика t будет достаточной для θ тогда и только тогда, когда

$$f_r(t, t_1, \ldots, t_{r-1} \mid \theta) = g(t \mid \theta)\, h_{r-1}(t_1, \ldots, t_{r-1} \mid t), \quad (17.67)$$

где h_{r-1} не зависит от θ для $r = 2, 3, \ldots, n$ и любого набора $t_1, \ldots, t_{r-1}$ *).

17.32 В том виде, в каком оно было дано, определение (17.67) не позволяет в каждом конкретном случае определить, существует ли достаточная статистика. Однако мы можем представить это определение в форме условия на функцию правдоподобия. Если последняя может быть записана в виде

$$L(x_1, \ldots, x_n \mid \theta) = g(t \mid \theta)\, k(x_1, \ldots, x_n), \quad (17.68)$$

где $g(t \mid \theta)$ является функцией от t и θ **) и k не зависит от θ, то легко видеть, что (17.67) можно вывести из (17.68). Действительно, умножим обе части (17.68) на дифференциал $dx_1 \ldots dx_n$ и сделаем преобразование

$$\begin{aligned} t &= t(x_1, \ldots, x_n) \\ t_i &= t_i(x_1, \ldots, x_n), \qquad i = 1, 2, \ldots, r-1, \\ t_i &= x_i, \qquad\qquad\quad i = r, \ldots, n-1 \end{aligned}$$

(здесь r — любое, а набор $t_1, \ldots, t_{r-1}$ произволен). Якобиан преобразования не содержит θ, и (17.68) может быть приведено к виду

$$g(t \mid \theta)\, dt\, l(t, t_1, \ldots, t_{r-1}, t_r, \ldots, t_{n-1}) \prod_{i=1}^{n-1} dt_i. \quad (17.69)$$

*) Определение достаточности обычно дается для $r = 2$, но нам кажется, что определение для любых r более естественно. Такое изменение в определении не налагает дополнительных ограничений на понятие достаточности.

**) Мы сохраняем обозначение $g(t \mid \theta)$, так как эта функция от t и от θ может всегда рассматриваться как маргинальное распределение t.

Проинтегрировав по лишним переменным $t_r, \ldots, t_{n-1}$, получим для совместного распределения $t, t_1, \ldots, t_{r-1}$ в точности выражение (17.67).

Следует заметить, что, интегрируя по $t_r, \ldots, t_{n-1}$, т. е. по $x_r, \ldots, x_{n-1}$, мы предполагали, что при этом не появляются члены, зависящие от θ. Это, конечно, так, если область определения распределения соответствующей переменной не зависит от θ; позже мы увидим (**17.40—1**), что интегрирование остается независимым от θ, даже если один или оба предела интегрирования зависят от θ.

Обратный результат также легко устанавливается. Положим в (17.67) $t_i = x_i$ и $r = n$ $(i = 1, 2, \ldots, n-1)$; тогда

$$f_r(t, x_1, x_2, \ldots, x_{n-1} \mid \theta) = g(t \mid \theta)\, h_{n-1}(x_1, \ldots, x_{n-1} \mid t). \quad (17.70)$$

Умножая обе части (17.70) на дифференциал $dt\, dx_1 \ldots dx_{n-1}$ и совершая преобразование

$$x_n = x_n(t, x_1, \ldots, x_{n-1}),$$
$$x_i = x_i, \quad i = 1, 2, \ldots, n-1,$$

немедленно получаем (17.68). Таким образом, (17.67) необходимо и достаточно для (17.68). Доказательство было проведено для непрерывных распределений. В дискретном случае оно получается проще, и читатель может провести его сам по той же схеме. Очень общее доказательство эквивалентности (17.67) и (17.68) было дано Халмошем и Сэвиджем (1949).

Мы рассматривали только случай $n \geqslant 2$. При $n = 1$ в качестве определения достаточности берется (17.68).

Достаточность и минимум дисперсии

17.33 Из необходимого и достаточного условия достаточности в форме (17.68) сразу вытекает одно интересное следствие. Беря от обеих частей логарифмы и дифференцируя, мы получаем

$$\frac{\partial \log L}{\partial \theta} = \frac{\partial \log g(t \mid \theta)}{\partial \theta}. \quad (17.71)$$

Сравнивая (17.71) и условие (17.27) существования МГД-оценки для $\tau(\theta)$, мы видим, что такая оценка существует, только когда имеется достаточная статистика. Действительно, (17.27) есть просто специальный случай (17.71), именно, когда

$$\frac{\partial \log g(t \mid \theta)}{\partial \theta} = A(\theta)\{t - \tau(\theta)\}. \quad (17.72)$$

Таким образом, хотя на первый взгляд требование достаточности может показаться более ограничительным условием, чем

достижение МГД, в действительности оно менее ограничительно. (17.71) выполняется всегда, когда выполняется (17.27), но даже если (17.27) не выполняется, может оказаться, что достаточная статистика существует.

Пример 17.15

Из рассуждений пункта **17.33** следует, что все найденные нами МГД-оценки (примеры 17.6, 17.8, 17.9, 17.10) будут достаточными статистиками.

Пример 17.16

Рассмотрим задачу оценивания параметра θ в распределении

$$dF(x) = dx/\theta, \quad 0 \leqslant x \leqslant \theta.$$

Функция правдоподобия (ФП) в этом случае равна

$$L(x \mid \theta) = \begin{cases} \theta^{-n}, & \text{если } x_{(n)} \leqslant \theta, \\ 0 & \text{в противном случае,} \end{cases}$$

поскольку все наблюдения, включая наибольшее из них $x_{(n)}$, не превосходят θ.

Можно также написать

$$L(x \mid \theta) = \theta^{-n} u(\theta - x_{(n)}), \tag{17.73}$$

где

$$u(z) = \begin{cases} 1, & z \geqslant 0, \\ 0, & z < 0. \end{cases}$$

Из (17.73) видно, что ФП представляется в виде произведения функции от $x_{(n)}$ и от θ

$$g(x_{(n)} \mid \theta) = \theta^{-n} u(\theta - x_{(n)})$$

на функцию от $x_1, \ldots, x_n$

$$k(x_1, \ldots, x_n) = 1.$$

Следовательно, согласно (17.68) $x_{(n)}$ есть достаточная статистика для θ.

17.34 Достаточная статистика t определяется соотношением (17.68) единственным образом с точностью до взаимно однозначных функций от t. Например, если $t = t(u)$, то (17.68) можно записать в виде

$$L(x \mid \theta) = g(t \mid \theta) \cdot |t'(u)| \frac{k(x)}{|t'(u)|},$$

и мы имеем

$$L(x \mid \theta) = g_1(u \mid \theta) k_1(x),$$

где $g_1(u|\theta)$ является плотностью для u, а k_1 не зависит от θ, т. е. статистика u достаточна для θ. Такая неоднозначность, однако, не создает на практике никаких трудностей, поскольку мы выбираем такие функции от t, которые являются состоятельными и даже обычно несмещенными оценками для θ. Достаточными статистиками в некоторых случаях могут быть и не взаимно однозначные функции от t (см. упражнение 23.31).

Если не учитывать неоднозначность этого вида, то достаточная статистика единственна. Действительно, если имеется две достаточные статистики t_1 и t_2, то из (17.67) получаем при $r = 2$

$$f_2(t_1, t_2 | \theta) = g_1(t_1 | \theta)\, h_1(t_2 | t_1) = g_2(t_2 | \theta)\, h_2(t_1 | t_2), \quad (17.74)$$

откуда следует функциональное соотношение

$$t_1 = k(t_2, \theta). \quad (17.75)$$

Но t_1 и t_2 являются функциями только от наблюдений и не зависят от θ, следовательно, из (17.75) вытекает, что t_1 и t_2 функционально связаны.

17.35 В **17.33** мы видели, что если существует МГД-оценка, то она является достаточной статистикой. Теперь мы докажем более общий результат, принадлежащий Рао (1945) и Блекуэллу (1947) и состоящий в том, что независимо от достижения какой-либо границы дисперсии несмещенная МД-оценка, если она существует, всегда является функцией от достаточной статистики.

Пусть t — достаточная статистика для θ, а t_1 — некоторая другая статистика с конечной дисперсией и математическим ожиданием

$$\mathbf{M}(t_1) = \tau(\theta). \quad (17.76)$$

Ввиду (17.68) можно записать (17.76) следующим образом:

$$\tau(\theta) = \int \ldots \int t_1 g(t | \theta)\, k(x_1, \ldots, x_n)\, dx_1, \ldots, dx_n. \quad (17.77)$$

Переходя в правой части (17.77) к новым переменным $t, x_2, x_3, \ldots, x_n$ и интегрируя по последним $n - 1$ из них, получим

$$\tau(\theta) = \int p(t)\, g(t | \theta)\, dt, \quad (17.78)$$

откуда следует, что существует функция $p(t)$ от достаточной статистики, являющаяся несмещенной оценкой для $\tau(\theta)$. Из вывода равенства (17.78) видно, что $p(t) = \mathbf{M}(t_1 | t)$. Далее,

$$\mathbf{D}t_1 = \mathbf{M}\{t_1 - \tau(\theta)\}^2 = \mathbf{M}\{[t_1 - p(t)] + [p(t) - \tau(\theta)]\}^2 =$$
$$= \mathbf{M}\{t_1 - p(t)\}^2 + \mathbf{M}\{p(t) - \tau(\theta)\}^2,$$

так как $\mathsf{M}\{t_1 - p(t)\}\{p(t) - \tau(\theta)\} = 0$ (что можно показать, беря условное математическое ожидание относительно t). Следовательно,

$$\mathsf{D}t_1 = \mathsf{M}\{t_1 - p(t)\}^2 + \mathsf{D}\{p(t)\} \geqslant \mathsf{D}\{p(t)\},$$

причем равенство имеет место в том и только в том случае, когда $t_1 \equiv p(t)$. Таким образом, $p(t) = \mathsf{M}(t_1 | t)$ имеет меньшую дисперсию, чем t_1. В **23.9—11** мы увидим, что обычно только одна функция от достаточной статистики имеет заданное математическое ожидание. В этом случае любой выбор t_1 приводит к одной и той же оценке $\mathsf{M}(t_1 | t)$, которая и будет единственной несмещенной МД-оценкой для $\tau(\theta)$.

Полученный результат не всегда легко использовать конструктивно, поскольку может оказаться, что $\mathsf{M}(t_1 | t)$ трудно вычислить (один важный класс случаев, когда оценка получается в явном виде, дает упражнение 17.24); однако он гарантирует нам, что несмещенная оценка с конечной дисперсией, являющаяся функцией от достаточной статистики, будет единственной МД-оценкой.

Распределения, обладающие достаточными статистиками

17.36 Теперь мы постараемся определить класс распределений, для параметра которых существует достаточная статистика. Сначала будет рассмотрен случай, когда область изменения варианты не зависит от параметра θ. Если t достаточна для θ в выборке из n независимых наблюдений, то из (17.71) следует, что

$$\frac{\partial \log L}{\partial \theta} = \sum_{j=1}^{n} \frac{\partial \log f(x_j | \theta)}{\partial \theta} = K(t, \theta), \tag{17.79}$$

где K есть некоторая функция от t и θ. Рассматривая (17.79) как уравнение относительно t, мы видим, что оно остается справедливым для любого значения θ, в частности, равного нулю. Отсюда ясно, что t может быть представлено в виде

$$t = M\left\{\sum_{j=1}^{n} k(x_j)\right\}, \tag{17.80}$$

где M и k — произвольные функции. Если $w = \sum k(x_j)$, то K будет функцией, обозначим ее $N(\theta, w)$, только от θ и w. Поэтому из (17.79) вытекает, если существуют написанные ниже производные, что

$$\frac{\partial^2 \log L}{\partial \theta \, \partial x_j} = \frac{\partial N}{\partial w} \frac{\partial w}{\partial x_j}. \tag{17.81}$$

Левая часть (17.81) зависит только от θ и от x_j, а $\partial w/\partial x_j$ — только от x_j. Отсюда следует, что $\partial N/\partial w$ зависит только от θ и x_j. Но производная $\partial N/\partial w$ симметрична относительно иксов и, следовательно, зависит лишь от θ. Поэтому, интегрируя ее по w, получим

$$N(\theta, w) = wp(\theta) + q(\theta),$$

где p и q — произвольные функции от θ. Таким образом, (17.79) принимает вид

$$\frac{\partial}{\partial\theta}\log L = \frac{\partial}{\partial\theta}\sum_j \log f(x_j\,|\,\theta) = p(\theta)\sum k(x_j) + q(\theta), \quad (17.82)$$

откуда

$$\frac{\partial}{\partial\theta}\log f(x\,|\,\theta) = p(\theta)\,k(x) + q(\theta)/n,$$

что дает необходимое условие существования достаточной статистики:

$$f(x\,|\,\theta) = \exp\{A(\theta)\,B(x) + C(x) + D(\theta)\}. \quad (17.83)$$

Этот результат, принадлежащий Дармуа (1935), Питмэну (1936) и Купмэну (1936), в точности совпадает с формулой (17.30), определяющей общий вид распределений экспоненциального семейства и выведенной как условие существования МГД-оценки для некоторой функции от θ.

Работа Л. Брауна (1964) посвящена условиям регулярности, достаточным для справедливости этого результата, и содержит ссылки на относящиеся к этому вопросу публикации.

Если выполняется (17.83), то в случае, когда область определения $f(x|\theta)$ не зависит от θ, функция правдоподобия будет достаточной статистикой для θ. Таким образом, в этом случае условие (17.83) является достаточным для того, чтобы распределение обладало достаточной статистикой.

Все перечисленные в примере 17.15 распределения имеют вид (17.83).

17.37 Итак при выполнении условий регулярности имеется взаимно однозначное соответствие между существованием достаточной статистики для θ и существованием МГД-оценки для некоторой функции от θ. Если выполняется (17.83), то существует достаточная статистика для θ и имеется ровно одна функция t от этой статистики (являющаяся сама достаточной статистикой), которая удовлетворяет (17.27) и оценивает, следовательно, некоторую функцию $\tau(\theta)$ с дисперсией, равной МГД. Для больших выборок, кроме того, любая функция от достаточной статистики оценивает свое математическое ожидание с

точностью, равной МГД (см. **17.23**). Наконец, для любого n любая функция от достаточной статистики будет для своего математического ожидания оценкой с минимальной достижимой дисперсией (см. **17.35**).

Достаточные статистики для нескольких параметров

17.38 Идеи предыдущих пунктов непосредственно обобщаются на случай распределений, зависящих от нескольких параметров $\theta_1, \ldots, \theta_k$. По существу, не появится также ничего нового, если мы будем вместо одномерных рассматривать многомерные распределения. Итак, пусть $\boldsymbol{x}_i$ — векторная варианта с $p(\geqslant 1)$ компонентами, $\boldsymbol{t}$ (или $\boldsymbol{t}_i$) — вектор, компонентами которого служат s статистик, и $\boldsymbol{\theta}$ — вектор параметров с k компонентами. Результаты пунктов **17.31** и **17.32** в этом случае почти не изменяются. Если можно написать соотношение

$$L(\boldsymbol{x} \mid \boldsymbol{\theta}) = g(\boldsymbol{t} \mid \boldsymbol{\theta})\, h(\boldsymbol{x}), \tag{17.84}$$

то совокупность компонент вектора $\boldsymbol{t}$ будет называться системой (совместно) достаточных статистик для $\boldsymbol{\theta}$. Свойство (17.67) выводится так же, как и раньше.

Число s компонент вектора $\boldsymbol{t}$ может быть больше, равно или меньше, чем число k компонент вектора параметров. Если $s = 1$, то мы имеем дело с *одномерной* достаточной статистикой. (Таким образом, термин «достаточная статистика», которым мы пользовались до сих пор, следовало бы заменить термином «одномерная достаточная статистика».) Полагая $\boldsymbol{t} = \boldsymbol{x}$, мы видим, что сами наблюдения *всегда* образуют систему достаточных статистик для $\boldsymbol{\theta}$ с $s = n$. Желая упростить задачу анализа данных, мы, естественно, хотим, чтобы s было как можно меньше. Но даже это требование не является достаточно ограничительным (см. упражнение 18.13). В главе 23 мы определим понятие минимальной системы достаточных статистик, которая будет функцией от всех других систем достаточных статистик.

Очевидно, что из совместной достаточности $\boldsymbol{t}$ для $\boldsymbol{\theta}$ не следует, что какая-нибудь отдельная компонента $\boldsymbol{t}$, скажем $t^{(1)}$, служит достаточной статистикой для θ_1. Так будет только в том случае, когда $g(\boldsymbol{t} \mid \boldsymbol{\theta})$ разлагается в произведение, одним из сомножителей которого является $g_1(t^{(1)} \mid \theta_1)$. Неверно и обратное: индивидуальная достаточность каждого из $t^{(i)}$ при известных остальных не обеспечивает совместной достаточности.

Для $k = 1$ результат пункта **17.35** сохраняется, если $\boldsymbol{t}$ — вектор, имеющий $s < n$ компонент (случай $s = n$ бессодержателен). При этом, чем меньше s, тем легче этот результат использовать.

Пример 17.17

Рассмотрим оценивание параметров μ и σ^2 распределения

$$dF(x) = \frac{1}{\sigma(2\pi)^{1/2}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} dx, \quad -\infty \leqslant x \leqslant \infty.$$

Поскольку

$$L(x \mid \mu, \sigma^2) = \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2\right\}, \quad (17.85)$$

а совместное распределение $\bar{x}$ и s^2 (пример 11.7) имеет вид

$$g(\bar{x}, s^2 \mid \mu, \sigma^2) \propto \frac{1}{\sigma} \exp\left\{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2\right\} \frac{1}{\sigma^{n-1}} s^{n-3} \exp\left\{-\frac{ns^2}{2\sigma^2}\right\},$$

то, учитывая соотношение $\sum (x-\mu)^2 = n\{s^2 + (\bar{x}-\mu)^2\}$, получим

$$L(x \mid \mu, \sigma^2) = g(\bar{x}, s^2 \mid \mu, \sigma^2)\, k(x).$$

Таким образом, $\bar{x}$ и s^2 совместно достаточны для μ и σ^2. Мы уже видели, что $\bar{x}$ достаточна для μ при известном σ^2 (примеры 17.6, 17.15), а $\frac{1}{n}\sum (x-\mu)^2$ достаточна для σ^2 при известном μ (примеры 17.10, 17.15). Из (17.85) легко видеть, что одна s^2 не будет достаточной для σ^2.

17.39 Основные результаты, относящиеся к достаточным статистикам, естественным образом обобщаются на случай k параметров. Условие (обобщающее (17.83)) того, что распределение обладает системой k статистик, совместно достаточных для k параметров, при аналогичных предположениях о непрерывности и существовании производных выглядит следующим образом:

$$f(x) = \exp\left\{\sum_{j=1}^{k} A_j(\theta) B_j(x) + C(x) + D(\theta)\right\}. \quad (17.86)$$

Оно было получено Дармуа (1935), Купмэном (1936) и Питмэном (1936). Более общие результаты, связанные с этим вопросом, получили Баранкин и Майтра (1963). Результат пункта **17.35** о свойствах функций от достаточных статистик как единственных оценках с минимальной дисперсией обобщается в следующей теореме Рао(1947): при одновременном оценивании r ($r \leqslant k$) функций τ_j от k параметров θ_s несмещенные оценки t_i, являющиеся функциями от минимальной системы из k достаточных статистик, имеют минимальные достижимые дисперсии. Нижние границы (не обязательно достижимые) дисперсий этих

функций определяются (если область изменения вариант не зависит от θ_s) соотношением

$$\mathsf{D}t_i \geqslant \sum_{j=1}^{k} \sum_{l=1}^{k} \frac{\partial \tau_i}{\partial \theta_j} \frac{\partial \tau_i}{\partial \theta_l} I_{jl}^{-1}, \quad i = 1, 2, \ldots, r, \qquad (17.87)$$

где I_{jl}^{-1} — элементы матрицы, обратной к *информационной матрице*

$$(I_{jl}) = \left\{ \mathsf{M}\left(\frac{1}{L} \frac{\partial L}{\partial \theta_j} \cdot \frac{1}{L} \frac{\partial L}{\partial \theta_l} \right) \right\}. \qquad (17.88)$$

Неравенство (17.87) обобщает неравенства (17.22) и (17.50), являющиеся его простейшими частными случаями. Подобно (17.22), (17.87) учитывает только члены порядка $1/n$.

Достаточность в случае зависящей от параметра области изменения варианты

17.40 Рассмотрим теперь ситуацию, в которой область изменения варианты зависит от θ. Мы опускаем тривиальный случай $n = 1$. Пусть сначала от θ зависит только нижняя граница области изменения. Тогда плотность распределения $f(x|\theta)$ будет определена на отрезке $a(\theta) \leqslant x \leqslant b$, где $a(\theta)$ — монотонная функция от параметра θ, изменяющегося в некотором невырожденном интервале.

Точно так же, как в примере 17.16, мы получим

$$L(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta)\, u(x_{(1)} - a(\theta)),$$

где $x_{(1)}$ — наименьшее значение в выборке, а

$$u(z) = \begin{cases} 1, & \text{если} \quad z \geqslant 0, \\ 0 & \text{в противном случае.} \end{cases}$$

Из вида функции правдоподобия ясно, что $x_{(1)}$ невозможно выделить в сомножитель, не зависящий от θ; поэтому одномерной достаточной статистикой, если такая существует, может быть только $x_{(1)}$. Но статистика $x_{(1)}$ может быть достаточной, только если $f(x_i|\theta)$ разлагается в произведение двух функций, одна из которых зависит от x_i, а другая — от θ, т. е. если

$$f(x|\theta) = g(x)/h(\theta). \qquad (17.89)$$

Это условие является необходимым и достаточным для того, чтобы функция правдоподобия имела вид

$$L(x|\theta) = \frac{u(x_{(1)} - a(\theta))}{\{h(\theta)\}^n} \cdot \prod_{i=1}^{n} g(x_i),$$

т. е. чтобы согласно (17.68) $x_{(1)}$ была достаточной.

Мы придем к такому же результату (с заменой $x_{(1)}$ на $x_{(n)}$) в случае варианты, меняющейся в пределах $a \leqslant x \leqslant b(\theta)$, где $b(\theta)$ — монотонная функция от θ: $x_{(n)}$ будет тогда и только тогда одномерной достаточной статистикой, когда выполняется (17.89).

При предположении (17.89) в упражнении 17.24 требуется, используя результат пункта **17.35**, найти явный вид единственной несмещенной МД-оценки для произвольной функции $\tau(\theta)$. Следует отметить, что в этой задаче параметр выбран так, что он совпадает с соответствующим пределом области изменения варианты.

17.41 Если оба предела изменения варианты зависят от θ, т. е. $a(\theta) \leqslant x \leqslant b(\theta)$, то мы имеем

$$L(x \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta)\, u(x_{(1)} - a(\theta))\, u(b(\theta) - x_{(n)}).$$

Рассуждения, аналогичные предыдущим, показывают, что необходимым и достаточным условием того, что крайние наблюдения $x_{(1)}$ и $x_{(n)}$ образуют систему достаточных статистик для θ, является (17.89). Посмотрим теперь, существует ли в данном случае одномерная достаточная статистика; заметим, что если такая статистика существует, то она является функцией от $x_{(1)}$ и от $x_{(n)}$.

По существу, мы должны решить вопрос о существовании одномерной статистики, которая показывала бы, чему равно произведение $u(x_{(1)} - a(\theta))\, u(b(\theta) - x_{(n)})$ в $L(x \mid \theta)$ — нулю или единице. Это произведение равно единице, если

$$x_{(1)} \geqslant a(\theta), \qquad b(\theta) \geqslant x_{(n)}. \tag{17.90}$$

Имеется четыре возможности:

(I) $a(\theta)$ и $b(\theta)$ — возрастающие функции от θ. Тогда (17.90) принимает вид

$$a^{-1}(x_{(1)}) \geqslant \theta \geqslant b^{-1}(x_{(n)}).$$

(II) $a(\theta)$ и $b(\theta)$ — убывающие функции от θ. Тогда (17.90) принимает вид

$$b^{-1}(x_{(n)}) \geqslant \theta \geqslant a^{-1}(x_{(1)}).$$

(III) $a(\theta)$ возрастает, $b(\theta)$ убывает. Тогда (17.90) принимает вид

$$a^{-1}(x_{(1)}) \geqslant \theta, \qquad b^{-1}(x_{(n)}) \geqslant \theta. \tag{17.91}$$

(IV) $a(\theta)$ убывает, $b(\theta)$ возрастает. Тогда (17.90) принимает вид

$$a^{-1}(x_{(1)}) \leqslant \theta, \qquad b^{-1}(x_{(n)}) \leqslant \theta. \tag{17.92}$$

В случаях (I) и (II) не существует одномерной достаточной статистики, так как необходимо знать и $x_{(1)}$ и $x_{(n)}$. Что касается случая (III), то (17.91) показывает, что функция $u(t_1 - \theta)$, где

$$t_1 = \min\{a^{-1}(x_{(1)}),\ b^{-1}(x_{(n)})\}, \tag{17.93}$$

эквивалентна произведению $u(x_{(1)} - a(\theta)) \cdot u(b(\theta) - x_{(n)})$, т. е. t_1 является одномерной достаточной статистикой для θ. Аналогично в случае (IV) одномерной достаточной статистикой будет

$$t_2 = \max\{a^{-1}(x_{(1)}),\ b^{-1}(x_{(n)})\}, \tag{17.94}$$

так как $u(\theta - t_2)$ эквивалентна произведению $u(x_{(1)} - a(\theta)) \times$ $\times\, u(b(\theta) - x_{(n)})$. Резюмируя, можно сказать, что если верхний предел есть монотонно убывающая функция нижнего и выполнено (17.89), то существует одномерная достаточная статистика, даваемая равенствами (17.93) или (17.94), соответственно тому, является $a(\theta)$ возрастающей или убывающей функцией от θ. Упражнение 20.13 дает распределение t_1, из которого легко получить и распределение t_2.

Изложенные результаты были получены Питмэном (1936) и Дэвисом (1951). Требование, чтобы θ принадлежало невырожденному интервалу, является существенным (см. пример 17.23).

Пример 17.18

Для прямоугольного распределения

$$dF = dx/(2\theta), \qquad -\theta \leqslant x \leqslant \theta,$$

имеет место случай (IV) пункта **17.41**. Одномерная достаточная статистика (17.94) принимает вид

$$t_2 = \max\{-x_{(1)},\ x_{(n)}\}$$

и совпадает ввиду неравенства $x_{(1)} \leqslant x_{(n)}$ со статистикой

$$t_2 = \max\{|x_{(1)}|,\ |x_{(n)}|\},$$

которая приемлема и с интуитивной точки зрения.

Пример 17.19

Распределение

$$dF(x) = \exp\{-(x - \alpha)\}\,dx, \qquad \alpha \leqslant x \leqslant \infty,$$

удовлетворяет условию (17.89), так как можно написать

$$f(x) = \exp(-x)/\exp(-\alpha).$$

Следовательно, наименьшее наблюдение $x_{(1)}$ является достаточной статистикой для нижнего предела α.

Пример 17.20

Распределение

$$dF(x) \propto \exp(-x\theta)\,dx, \quad 0 \leqslant x \leqslant \theta,$$

очевидно, не может быть приведено к виду (17.89); следовательно, при $n \geqslant 2$ для θ не существует одномерной достаточной статистики.

Пример 17.21

Рассмотрим распределение

$$dF(x) = dx/(\beta - \alpha), \quad \alpha \leqslant x \leqslant \beta,$$

зависящее от двух параметров. Ясно, что при фиксированном α $x_{(n)}$ достаточна для β, а при фиксированном β $x_{(1)}$ достаточна для α. Кроме того, $x_{(1)}$ и $x_{(2)}$ совместно достаточны для α и β. Это следует из того, что совместное распределение $x_{(1)}$ и $x_{(n)}$, даваемое формулой (14.2) с $r = 1$, $s = n$, имеет вид

$$g(x_{(1)}, x_{(n)}) = n(n-1)(x_{(n)} - x_{(1)})^{n-2}/(\beta - \alpha)^n,$$

и поэтому можно написать

$$L(x \mid \alpha, \beta) = (\beta - \alpha)^{-n} = g(x_{(1)}, x_{(n)})\, k(x).$$

Пример 17.22

Прямоугольному распределению

$$dF(x) = dx/\theta, \quad k\theta \leqslant x \leqslant (k+1)\theta, \quad k > 0,$$

соответствует случай (I) пункта **17.41**. Одномерной достаточной статистики здесь не существует, но пара $(x_{(1)}, x_{(n)})$ образует достаточную систему, так как (17.89) выполняется.

Пример 17.23

В случае прямоугольного распределения

$$dF = dx, \quad \theta \leqslant x \leqslant \theta + 1, \quad \theta = 0, 1, 2, \ldots,$$

когда θ может принимать только целочисленные значения, верхний предел не является монотонно убывающей функцией нижнего предела. Однако очевидно, что *любое* наблюдение x_i из выборки объема n будет одномерной достаточной статистикой для θ. В действительности $[x_i]$ оценивает θ даже с нулевой дисперсией. Если ограничение, что θ принимает только целочисленные значения, исключить, то, в соответствии с **17.41**, одномерной достаточной статистики существовать не будет.

17.42 На этом мы заканчиваем обсуждение основных идей теории оценивания. Дальнейшее развитие теории достаточных статистик будет дано в главах 23, 24. А пока мы продолжим изучение теории оценивания с другой точки зрения — мы исследуем свойства оценок, получаемых методом максимального правдоподобия.

УПРАЖНЕНИЯ

17.1 Показать, что для выборок из распределения

$$dF = \frac{1}{\Gamma(p)\,\theta^p}\, x^{p-1} e^{-x/\theta}\, dx, \qquad p > 0, \qquad 0 \leqslant x \leqslant \infty,$$

$\bar{x}/p$ является при фиксированном p МГД-оценкой для θ с дисперсией $\theta^2/(np)$, в то время как при $\theta = 1$ статистика $\frac{1}{n}\sum_{i=1}^{n} \log x_i$ есть МГД-оценка для $\frac{\partial}{\partial p} \log \Gamma(p)$ с дисперсией $\left\{ \frac{\partial^2 \log \Gamma(p)}{\partial p^2} \right\} \Big/ n$.

17.2 Случайная величина x имеет функцию плотности

$$f(x \mid \theta) = \theta f_1(x) + (1 - \theta) f_2(x),$$

где $0 < \theta < 1$, а f_1, f_2 — полностью известные ф. п., области определения которых не зависят от θ. Показать, что МГД при оценивании θ по выборке объема n равна

$$\frac{\theta(1-\theta)}{n} \left\{ 1 - \int_{-\infty}^{\infty} \frac{f_1(x)\, f_2(x)}{f(x \mid \theta)}\, dx \right\}^{-1}$$

и сводится к выражению для биномиального случая (пример 17.9), если области определения f_1 и f_2 не перекрываются.

(Хилл, 1963b.)

17.3 Вычислить $J_{22} = \mathbf{M}\left\{ \left(\frac{1}{L} \frac{\partial^2 L}{\partial \theta^2} \right)^2 \right\}$ для биномиального распределения в примере 17.11 и использовать полученный результат для точного вычисления (17.45). Сравнить результат с точной дисперсией t при $\theta = 1/2$, приведенной в этом примере.

17.4 Записывая (17.27) в виде

$$(t - \tau) = \frac{\tau'(\theta)}{J_{11}} \cdot \frac{L'}{L}$$

и обозначая характеристическую функцию $t - \tau(\theta)$ через $\varphi(z)$ $(z = iu)$, показать, что если t есть МГД-оценка, то справедливо соотношение

$$\frac{\partial \varphi(z)}{\partial z} = \frac{\tau'(\theta)}{J_{11}} \left\{ \frac{\partial \varphi(z)}{\partial \theta} + z\tau'(\theta)\, \varphi(z) \right\},$$

а семиинварианты t удовлетворяют равенствам

$$\varkappa_{r+1} = \frac{\partial \tau}{\partial \theta} \cdot \frac{\partial \varkappa_r}{\partial \theta} \Big/ J_{11}, \qquad r = 2, 3, \ldots \qquad \text{(A)}$$

Вывести отсюда, что ковариация между оценкой t и несмещенной оценкой ее r-го семиинварианта равна $(r+1)$-му семиинварианту t.

Доказать неравенство (17.50) для оценки функции от нескольких параметров и показать, что в случае достижения границы формула (А) останется справедливой.

(Бхаттачария, 1946, 1947.)

17.5 Показать, что в случае оценивания параметра θ логистического распределения

$$f(x) = e^{-(x-\theta)} \{1 + e^{-(x-\theta)}\}^{-2}$$

МГД равна $3/n$, тогда как точная дисперсия выборочного среднего равна $\pi^2/(3n)$, а асимптотическая дисперсия медианы равна $4/n$.

17.6 Показать, что для распределения

$$dF = \frac{1}{\sigma (2\pi)^{1/2}} \exp\left(-\frac{1}{2}\frac{x^2}{\sigma^2}\right) dx, \qquad -\infty \leqslant x \leqslant \infty,$$

статистики

$$t_1 = \left(\frac{1}{2}\sum_{i=1}^{n} x_i^2\right)^{1/2} \Gamma\left(\frac{1}{2}n\right) \Big/ \Gamma\left\{\frac{1}{2}(n+1)\right\}$$

и

$$t_2 = \left\{\frac{1}{2}\sum_{i=1}^{n} (x_i - \bar{x})^2\right\}^{1/2} \Gamma\left\{\frac{1}{2}(n-1)\right\} \Big/ \Gamma\left(\frac{1}{2}n\right)$$

являются несмещенными оценками параметра σ. Показать также, что (17.53) вообще дает бо́льшую границу, чем граница $\sigma^2/(2n)$, даваемая (17.24); проверить, рассматривая случай $n = 2$, что даже эта граница не достигается для оценки t_1 при малых n.

(Чепмэн и Роббинс, 1951.)

17.7 Показать, что в случае распределения

$$dF = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x-\mu)^2\right\} dx$$

$(\bar{x}^2 - 1/n - \mu^2)$ является линейной функцией от $\frac{1}{L}\frac{\partial L}{\partial \mu}$ и $\frac{1}{L}\frac{\partial^2 L}{\partial \mu^2}$ и, следовательно, согласно (17.42) $\bar{x}^2 - 1/n$ будет несмещенной оценкой для μ^2 с минимальной достижимой дисперсией.

17.8 Исходя непосредственно из вида функций правдоподобия, показать, что в случаях, рассмотренных в примерах 17.8 и 17.9, достаточные статистики существуют.

17.9 Используя (17.86), показать, что в случае трехпараметрического распределения

$$dF(x) = \frac{1}{\Gamma(p)} \exp\left\{-\left(\frac{x-\alpha}{\sigma}\right)\right\} \left(\frac{x-\alpha}{\sigma}\right)^{p-1} \frac{dx}{\sigma}, \quad p, \sigma > 0, \quad \alpha \leqslant x \leqslant \infty,$$

а) существует достаточная статистика для p при известных α и σ; б) существует достаточная статистика для σ при известных p и α; в) существуют совместно достаточные статистики для p и σ при известном α; г) имеется достаточная статистика для α при известном σ и $p = 1$.

17.10 Доказать неравенство (17.87) для дисперсии оценки, входящей в систему оценок для нескольких функций от параметров.

17.11 Показать, что если t_1 есть несмещенная МД-оценка для θ, а t_2 — другая несмещенная оценка, то ковариация t_1 и $(t_2 - t_1)$ равна нулю. Вывести отсюда, что отклонение $(t_2 - \theta)$ можно разложить на две компоненты, одна из которых соответствует отклонению $(t_1 - \theta)$, а другая — неэффективности оценки.

(См. Фишер, 1925.)

17.12 Пусть в биномиальном распределении примера 17.9 a и b принимают целые неотрицательные значения. Положим $\tau_{ab} = \theta^a (1-\theta)^b$. Показать, что если $a, b \geqslant 0$ и $a + b \leqslant n$, то τ_{ab} имеет несмещенную оценку *) $r^{(a)}(n-r)^{(b)}/n^{(a+b)}$, и что в других случаях несмещенной оценки для τ_{ab} не существует.

17.13 Пусть дана выборка объема n из распределения вида (17.89) с областью определения $a \leqslant x \leqslant \theta$. Показать, что достаточная статистика $x_{(n)}$ не будет несмещенной оценкой для θ. Используя метод пункта **17.10**, показать, что $t' = 2x_{(n)} - x_{(n-1)}$ не смещена до членов порядка n^{-1} и имеет такой же, до членов порядка n^{-2}, средний квадрат ошибки, как и $x_{(n)}$, именно $\frac{2}{n^2}\left\{\frac{h(\theta)}{h'(\theta)}\right\}^2$. Показать далее, что $t'' = 3(x_{(n)} - x_{(n-1)}) + x_{(n-2)}$.

(См. Робсон и Уитлок, 1964.)

17.14 Пусть дана выборка объема n из распределения с функцией плотности (17.86), область определения которой не зависит от параметров. Показать, что статистики

$$t_j = \sum_{i=1}^{n} B_j(x_i), \qquad j = 1, 2, \ldots, k,$$

образуют систему из k совместно достаточных статистик для параметров $\theta_1, \ldots, \theta_k$ и что их совместная функция плотности дается выражением

$$g(t_1, t_2, \ldots, t_k \mid \theta) = \exp\{nD(\theta)\}\, h(t_1, t_2, \ldots, t_k) \exp\left\{\sum_{j=1}^{k} A_j(\theta)\, t_j\right\},$$

т. е. сама имеет вид (17.86).

17.15 Использовать результат упражнения 17.14 для вывода распределения $\bar{x}$ в примере 17.6.

17.16 Показать, что статистика $\frac{1}{(n+1)}\sum (x - \bar{x})^2$, отличающаяся множителем от выборочной дисперсии, представляет собой оценку дисперсии нормальной совокупности с наименьшим средним квадратом ошибки.

17.17 Использовать метод пункта **17.10** для исключения смещения в выборочной дисперсии при оценивании генеральной дисперсии. (Ср. с результатом примера 17.3.)

(Кэнуй, 1956.)

17.18 Показать, что при использовании метода, изложенного в пункте **17.10**, дисперсия t_n' совпадает с дисперсией t_n до членов порядка $1/n$.

(Кэнуй, 1956.)

17.19 Использовать метод пункта **17.10** для получения поправки на смещение при оценивании квадрата генерального среднего квадратом выборочного среднего.

---

*) Здесь используется обозначение $n^{(a)} = n(n-1)\ldots(n-a+1)$ (*Прим. ред.*)

17.20 Для (положительно определенной) матрицы ковариаций произведение любого диагонального элемента на соответствующий диагональный элемент обратной матрицы не может быть меньше единицы. Вывести отсюда, что если в (17.87) мы имеем $r = k$ и $\tau_i = \theta_i$ (для всех i), то получающаяся для θ_i граница будет не меньше границы, даваемой (17.24). Объяснить, почему получается такой результат.

(С. Р. Рао, 1952.)

17.21 Неравенство (17.22), определяющее МГД, остается верным для распределения $f(x|\theta)$, у которого границы области определения (a, b) зависят от θ, при условии, что выполняется (17.18). Показать, что это так, если

$$f(a \mid \theta) = f(b \mid \theta) = 0,$$

и что если, кроме того,

$$\left[\frac{\partial f(x \mid \theta)}{\partial \theta}\right]_{x=a} = \left[\frac{\partial f(x \mid \theta)}{\partial \theta}\right]_{x=b} = 0,$$

то (17.19) тоже остается верным, и мы можем представить МГД в форме (17.23).

17.22 Используя результат упражнения 17.21, показать, что при оценивании параметра θ в распределении

$$dF(x) = \frac{1}{\Gamma(p)}(x-\theta)^{p-1}\exp\{-(x-\theta)\}\,dx, \quad \theta \leqslant x \leqslant \infty, \quad p > 2,$$

МГД дается неравенством (17.22) и равна $(p-2)/n$, однако является недостижимой, так как для θ не существует одномерной достаточной статистики.

17.23 Пусть случайная величина x имеет распределение $f(x|\theta)$ и изменяется в интервале (a, b) (который может зависеть от θ). Показать, что если $\mathbf{M}\{t(x, \theta)\} = \tau(\theta)$ и f вместе с $\partial f/\partial x$ обращается в нуль в точках a и b, то

$$\mathrm{D}t \geqslant \left[\mathbf{M}\left(\frac{\partial t}{\partial x}\right)\right]^2 \Big/ \mathbf{M}\left\{\left(\frac{\partial \log f}{\partial x}\right)^2\right\}$$

(ср. с выводом (17.22) в упражнении 17.21).

(Б. Рао (1958b). Имеет место также аналог (17.45) с заменой $\tau^{(r)}$ на $\mathbf{M}\left(\frac{\partial^r t}{\partial x^r}\right)$ и $\frac{L^{(r)}}{L}$ в J_{rp} на $\frac{1}{f}\frac{\partial^r f}{\partial x^r}$ — см. Санкаран (1964).)

17.24 Пусть

$$f(x) = g(x)/h(\theta), \qquad \theta \leqslant x \leqslant b.$$

Показать, что если функция $t(x)$ от одного наблюдения является несмещенной оценкой для $\tau(\theta)$, то

$$-t(x) = \{\tau(x)\,h'(x) + \tau'(x)\,h(x)\}/g(x),$$

Используя **17.35**, вывести отсюда, что единственная МД-оценка для $\tau(\theta)$ по выборке объема n дается выражением

$$p(x_{(1)}) = \tau(x_{(1)}) - \frac{\tau'(x_{(1)})}{n}\frac{h(x_{(1)})}{g(x_{(1)})}$$

и аналогично для $a \leqslant x \leqslant \theta$

$$p(x_{(n)}) = \tau(x_{(n)}) + \frac{\tau'(x_{(n)})}{n} \frac{h(x_{(n)})}{g(x_{(n)})}.$$

В качестве иллюстрации рассмотреть распределения

$$f(x) = 1/\theta, \qquad 0 \leqslant x \leqslant \theta,$$

и

$$f(x) = \exp\{-(x-\theta)\}, \qquad \theta \leqslant x \leqslant \infty.$$

(Тейт, 1959)

17.25 Пусть две статистики (t_1, t_2) совместно достаточны для двух параметров (θ_1, θ_2) и t_1 достаточна для θ_1 при известном θ_2. Показать, что условное распределение t_2 при данном t_1 не зависит от θ_1. Рассмотреть в качестве иллюстрации k независимых биномиальных выборок, имеющих объемы n_i $(i = 1, 2, \ldots, k)$ и параметры θ_i, связанные соотношением

$$\log\left(\frac{\theta_i}{1-\theta_i}\right) = \alpha + \beta x_i,$$

и показать, что если y_i есть число «успехов» в i-й выборке, то условное распределение суммы $\sum_i x_i y_i$ при данной сумме $\sum_i y_i$ не зависит от α.

(Д. Кокс, 1958а)

17.26 Если нулевой исход в пуассоновском распределении не может быть наблюдаем, то оно называется усеченным пуассоновским. Показать, что единственная несмещенная оценка для $1 - e^{-\theta}$, построенная по одному наблюдению x $(x = 1, 2, \ldots)$ над усеченным пуассоновским распределением $e^{-\theta}\theta^x/x!$, равна 0 при x нечетном и 1 при x четном.

17.27 Показать, подобно тому как это было сделано в примере 17.13, что эффективность рассмотренной в **10.14** оценки для σ, основанной на средней разности, в случае нормальных выборок равна 0,978.

ГЛАВА 18

ОЦЕНИВАНИЕ: МЕТОД МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

18.1 В **8.6—10** мы уже познакомились с принципом максимального правдоподобия (МП) в его общей формулировке. Данная глава посвящена приложениям этого принципа к задачам оценивания и изучению его свойств как метода оценивания. При этом мы в основном будем иметь дело с выборками из n независимых наблюдений из одного и того же распределения. Совместное распределение наблюдений, рассматриваемое как функция неизвестного параметра θ, будет называться **функцией правдоподобия** (ФП) выборки

$$L(x\,|\,\theta) = f(x_1\,|\,\theta) f(x_2|\,\theta) \ldots f(x_n\,|\,\theta), \tag{18.1}$$

где $f(x|\theta)$ обозначает плотность распределения, если оно непрерывно, либо вероятность значения x, если распределение дискретно, причем распределение может быть как одномерным, так и многомерным. Согласно принципу МП, который после появления работы Фишера (1921а) стал широко использоваться в статистической теории, в качестве оценки для θ следует взять то значение $\hat{\theta}$ из области допустимых значений параметра θ, для которого ФП принимает наибольшее возможное значение. Другими словами, надо найти такое $\hat{\theta}$, для которого при любом допустимом значении θ выполняется неравенство

$$L(x\,|\,\hat{\theta}) \geqslant L(x\,|\,\theta). \tag{18.2}$$

Далее будет предполагаться, что θ может принимать все действительные значения из некоторого интервала (который может быть бесконечным в одну или в обе стороны).

18.2 В одном довольно общем случае вид оценки максимального правдоподобия (МП-оценки) определяется сравнительно просто. Если ФП дважды дифференцируема в своей области определения, то ее стационарные значения (если они существуют) даются корнями уравнения

$$L'(x|\,\theta) = \frac{\partial L(x\,|\,\theta)}{\partial \theta} = 0. \tag{18.3}$$

Достаточным (но не необходимым) условием того, чтобы некоторое стационарное значение $\tilde{\theta}$ было локальным максимумом, является неравенство

$$L''(x \mid \tilde{\theta}) < 0. \tag{18.4}$$

Если таким путем мы найдем все локальные максимумы ФП и затем выберем среди них наибольший (если их несколько), то мы получим решение (или решения) (18.2) при условии, что ФП не имеет краевого максимума в граничных точках области изменения θ.

18.3 На практике часто проще иметь дело с логарифмом ФП, чем с самой функцией. При выполнении условий предыдущего пункта ФП и ее логарифм будут иметь максимумы в одних и тех же точках, так как

$$\frac{\partial}{\partial\theta} \log L = \frac{L'}{L}$$

и $L > 0$. Поэтому в тех случаях, когда это проще, ищут решения уравнения

$$(\log L)' = 0, \tag{18.5}$$

для которых

$$(\log L)'' < 0. \tag{18.6}$$

(18.5) часто называется уравнением правдоподобия.

Максимальное правдоподобие и достаточность

18.4 Легко видеть, что в случае, когда для θ существует одномерная достаточная статистика, МП-оценка должна быть функцией от этой статистики. Действительно, если t достаточна для θ, то ФП разлагается в произведение двух сомножителей (см. (17.84)):

$$L(x \mid \theta) = g(t \mid \theta)\, h(x). \tag{18.7}$$

Поскольку второй сомножитель в правой части (18.7) не зависит от θ, то выбор $\hat{\theta}$, максимизирующего $L(x \mid \theta)$, эквивалентен выбору $\hat{\theta}$, максимизирующего $g(t \mid \theta)$, т. е. $\hat{\theta}$ есть функция только от t.

18.5 Если для $\tau(\theta)$ существует МГД-оценка t, а уравнение правдоподобия (18.5) имеет решение $\hat{\theta}$, то $t = \tau(\hat{\theta})$, решение $\hat{\theta}$ единственно и обращает в максимум ФП. Действительно, в **17.33** было показано, что в случае существования одномерной достаточной статистики ФП такова, что некоторая функция от θ может иметь МГД-оценку. Таким образом, как и в (17.27), можно написать

$$(\log L)' = A(\theta)\{t - \tau(\theta)\}, \tag{18.8}$$

откуда следует, что решения (18.5) имеют вид

$$t = \tau(\hat{\theta}). \tag{18.9}$$

Дифференцируя (18.8) еще раз, получим

$$(\log L)'' = A'(\theta)\{t - \tau(\theta)\} - A(\theta)\tau'(\theta). \tag{18.10}$$

Но так как согласно (17.29)

$$\tau'(\theta)/A(\theta) = \mathrm{D}t,$$

то последний член в (18.10) может быть представлен в виде

$$-A(\theta)\tau'(\theta) = -\{A(\theta)\}^2\,\mathrm{D}t. \tag{18.11}$$

Кроме того, при $\theta = \hat{\theta}$ первый член в правой части (18.10) в силу (18.9) равен нулю. Следовательно, используя (18.11), мы получаем из (18.10)

$$(\log L)''_{\theta} = -\{A(\hat{\theta})\}^2\,\mathrm{D}t < 0. \tag{18.12}$$

Из (18.12) следует, что каждое решение (18.5) является максимумом ФП. Но в регулярном случае между последовательными максимумами должен находиться минимум. Поскольку же не имеется ни одного минимума, то не может быть больше одного максимума. Это, с другой стороны, следует также из единственности МГД-оценки.

(18.9) показывает, что если (несмещенная) МГД-оценка существует, то она дается методом МП.

18.6 Утверждение о единственности МП-оценки при наличии одномерной достаточной статистики распространяется на случай, когда область определения $f(x|\theta)$ зависит от θ. Однако доказательство несколько меняется. В **17.40—1** было показано, что необходимым условием существования достаточной статистики является возможность разложения

$$f(x|\theta) = g(x)/h(\theta). \tag{18.13}$$

В этом случае ФП имеет вид

$$L(x|\theta) = \prod_{i=1}^{n} g(x_i)/\{h(\theta)\}^n \tag{18.14}$$

и принимает наибольшее значение, когда $h(\theta)$ принимает наименьшее. Из (18.13) получаем

$$1 = \int f(x|\theta)\,dx = \int g(x)\,dx/h(\theta),$$

где интегрирование производится по всей области изменения x. Следовательно,

$$h(\theta) = \int g(x)\, dx, \qquad (18.15)$$

откуда вытекает, что минимизация $h(\theta)$ эквивалентна минимизации интеграла в правой части (18.15) (один или оба предела интегрирования которого зависят от θ).

Вспомним теперь (**17.40—1**), что одномерная достаточная статистика для θ существует, когда либо только один предел зависит от θ, либо верхний предел является монотонно убывающей функцией нижнего. В обоих случаях значение (18.15) будет монотонной функцией от области интегрирования, достигающей единственного краевого минимума для наименьшей возможной при данных наблюдениях области. МП-оценка θ, полученная в результате минимизации области интегрирования в (18.15), единственна, и краевой максимум ФП (18.14) равен $L(x|\hat{\theta})$.

Результаты этого и предыдущего пунктов были впервые получены Хузурбазаром (1948) (который, однако, в «регулярном» случае **18.5** применил другой метод).

18.7 Таким образом, мы показали, что если для параметра θ существует одномерная достаточная статистика, то МП-оценка $\hat{\theta}$ для θ есть функция только от t. Далее, в этом случае $\hat{\theta}$ единственна и ФП имеет только один максимум. Этот максимум будет стационарным значением (при выполнении условий регулярности) либо краевым максимумом, если, соответственно, область определения не зависит от θ или зависит.

18.8 Из наших результатов следует, что МП-оценки, являющиеся взаимно однозначными функциями достаточных статистик, обладают всеми оптимальными свойствами последних. Например, нам достаточно получить решение уравнения правдоподобия и найти функцию от этого решения, которая была бы несмещенной оценкой для параметра. Согласно **17.35** эта функция и будет единственной МД-оценкой, для которой достигается (если это возможно) МГД (17.22).

Достаточные статистики, полученные в примерах 17.8—17.10, 17.16, 17.18 и 17.19, легко находятся с помощью метода МП.

Пример 18.1

Оценим параметр θ в распределении $dF(x) = dx/\theta$, $0 \leqslant x \leqslant \theta$. Из вида ФП, приведенной в примере 17.16, следует, что $\hat{\theta} = x_{(n)}$ является достаточной статистикой и ФП имеет в этой точке недифференцируемый максимум.

Очевидно, $\hat{\theta}$ не является несмещенной оценкой для θ. Легко видеть, что несмещенной будет оценка

$$t = (n+1)\,x_{(n)}/n.$$

Пример 18.2

Оценим среднее θ нормального распределения с известной дисперсией. В примере 17.6 мы получили

$$(\log L)' = \frac{n}{\sigma^2}(\bar{x} - \theta).$$

Приравнивая это выражение нулю, находим

$$\hat{\theta} = \bar{x}.$$

В этом случае $\hat{\theta}$ есть несмещенная оценка для θ.

Общий случай

18.9 Если для θ не существует одномерной достаточной статистики, то ФП уже не обязательно будет иметь только один максимум (см. упражнения 18.17, 18.33). В этом случае для выбора МП-оценки надо воспользоваться условием (18.2). Далее будут рассматриваться свойства полученных таким путем оценок. Мы увидим, что при достаточно широких условиях МП-оценки состоятельны, а при выполнении условий регулярности (наиболее важным из которых является независимость области определения $f(x|\theta)$ от θ), кроме того, асимптотически нормальны и эффективны. Однако указанные свойства относятся к большим выборкам, а поэтому, как бы ни были важны эти свойства, надо всегда помнить, что они не могут быть такими же убедительными аргументами в пользу МП-оценок, какими были рассмотренные нами свойства, связанные с достаточными статистиками. Возможно, вообще не следует ожидать от какого-то одного метода, чтобы он давал «наилучшие» результаты при любых обстоятельствах и для всех объемов выборок. Как бы то ни было, остается фактом то, что вне связи с достаточными статистиками МП-оценки обладают только асимптотическими оптимальными свойствами.

Пример 18.3

В качестве примера общей ситуации рассмотрим оценивание параметра корреляции ρ по выборке объема n из нормированного двумерного нормального распределения

$$dF = \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right\} dx\,dy,$$
$$-\infty \leqslant x,\ y \leqslant \infty, \quad |\rho| < 1.$$

Находим

$$\log L = -n \log (2\pi) - \frac{1}{2} n \log (1-\rho^2) - \\ - \frac{1}{2(1-\rho^2)} \left(\sum x^2 - 2\rho \sum xy + \sum y^2 \right),$$

откуда для $\frac{\partial \log L}{\partial \rho} = 0$ получаем

$$\frac{n\rho}{1-\rho^2} - \frac{\rho}{(1-\rho^2)^2} \left(\sum x^2 - 2\rho \sum xy + \sum y^2 \right) + \frac{1}{1-\rho^2} \sum xy = 0.$$

Это приводится к кубическому уравнению

$$\rho (1-\rho^2) + (1+\rho^2) \frac{1}{n} \sum xy - \rho \left(\frac{1}{n} \sum x^2 + \frac{1}{n} \sum y^2 \right) = 0.$$

Последнее имеет три корня, два из которых могут быть комплексными. В случае, когда все три корня действительны и соответствуют допустимым значениям ρ, мы согласно (18.2) выбираем в качестве МП-оценки тот из них, для которого ФП принимает наибольшее значение.

Если записать кубическое уравнение в виде

$$\lambda^3 + p\lambda + q = 0,$$

где

$$\lambda = \rho - \frac{1}{3n} \sum xy$$

и

$$p = \frac{1}{n} \sum x^2 + \frac{1}{n} \sum y^2 - \frac{1}{3} \left(\frac{1}{n} \sum xy \right)^2 - 1, \qquad (18.16)$$

то условием того, что оно имеет единственный корень, является неравенство

$$4p^3 + 27q^2 > 0.$$

Это неравенство наверняка выполнено при $p > 0$. Так как согласно **10.3** и **10.9** выборочные моменты в (18.16) являются состоятельными оценками соответствующих генеральных моментов, то p сходится по вероятности к положительной величине

$$\left(1 + 1 - \frac{1}{3}\rho^2 - 1\right) = 1 - \frac{1}{3}\rho^2 > 0.$$

Таким образом, асимптотически уравнение правдоподобия имеет только один действительный корень, который и будет МП-оценкой, поскольку комплексные корни не являются допустимыми значениями.

Состоятельность оценок максимального правдоподобия

18.10 Теперь мы покажем, что при очень общих условиях МП-оценки состоятельны.

Рассмотрим, как и в (18.2), выборку из n независимых наблюдений из распределения $f(x|\theta)$ и для каждого n выберем МП-оценку θ. Тогда для любого допустимого значения θ будем иметь *)

$$\log L(x|\hat{\theta}) \geqslant \log L(x|\theta). \tag{18.17}$$

Пусть θ_0 обозначает истинное значение θ, а $\mathbf{M}_0$ — соответствующее математическое ожидание. Определим случайную величину $L(x|\theta)/L(x|\theta_0)$. В силу того, что геометрическое среднее невырожденного распределения не превосходит арифметического среднего, для всех $\theta^* \neq \theta_0$ будет выполняться неравенство

$$\mathbf{M}_0\left\{\log \frac{L(x|\theta^*)}{L(x|\theta_0)}\right\} < \log \mathbf{M}_0\left\{\frac{L(x|\theta^*)}{L(x|\theta_0)}\right\}. \tag{18.18}$$

Поскольку математическое ожидание в правой части равно

$$\int \ldots \int \frac{L(x|\theta^*)}{L(x|\theta_0)} L(x|\theta_0)\, dx_1 \ldots dx_n = 1,$$

то (18.18) записывается в виде

$$\mathbf{M}_0\left\{\log \frac{L(x|\theta^*)}{L(x|\theta_0)}\right\} < 0.$$

Вводя множитель $1/n$, получаем

$$\mathbf{M}_0\left\{\frac{1}{n}\log L\{x|\theta^*)\right\} < \mathbf{M}_0\left\{\frac{1}{n}\log L(x|\theta_0)\right\} \tag{18.19}$$

при условии, что математическое ожидание в правой части (18.19) существует. Это условие обычно бывает выполнено.

Далее, для любого θ выражение

$$\frac{1}{n}\log L(x|\theta) = \frac{1}{n}\sum_{i=1}^{n} \log f(x_i|\theta)$$

является средним n независимых одинаково распределенных случайных величин с математическим ожиданием

$$\mathbf{M}_0\left\{\log f(x|\theta)\right\} = \mathbf{M}_0\left\{\frac{1}{n}\log L(x|\theta)\right\}.$$

Следовательно, согласно усиленному закону больших чисел (7.25) $\frac{1}{n}\log L(x|\theta)$ при возрастании n сходится с вероятностью

*) Знак равенства в (18.17) может оказаться причиной неоднозначности в выборе последовательности значений $\hat{\theta}$. См. **18.11** и **18.13** ниже.

единица к своему математическому ожиданию. Отсюда и из (18.19) следует, что при $n \to \infty$ с вероятностью единица

$$\frac{1}{n} \log L(x \mid \theta^*) < \frac{1}{n} \log L(x \mid \theta_0),$$

или

$$\lim_{n \to \infty} \mathsf{P}\{\log L(x \mid \theta^*) < \log L(x \mid \theta_0)\} = 1, \qquad \theta^* \neq \theta_0. \quad (18.20)$$

С другой стороны, (18.17) при $\theta = \theta_0$ дает

$$\log L(x \mid \hat{\theta}) \geqslant \log L(x \mid \theta_0). \quad (18.21)$$

(18.20) и (18.21) показывают, что при $n \to \infty$ значение $L(x \mid \hat{\theta})$ не может отличаться от $L(x \mid \theta_0)$. Поэтому, если $L(x \mid \theta)$ является взаимно однозначной функцией от θ, то

$$\mathsf{P}\{\lim_{n \to \infty} \hat{\theta} = \theta_0\} = 1. \quad (18.22)$$

Мы привели эвристический вариант строгого доказательства состоятельности МП-оценки, данного Вальдом (1949). (Строгое доказательство требует некоторых дополнительных условий.)

18.11 Мы показали, что любая последовательность оценок, удовлетворяющая (8.2), состоятельна. Этот результат был усилен Хузурбазаром (1948), который доказал, что если выполнены условия регулярности, то при достаточно больших n существует единственная состоятельная МП-оценка.

Пусть ФП дважды дифференцируема. Тогда из сходимости $\hat{\theta}$ к θ_0 по вероятности следует, что

$$\frac{1}{n}\left[\frac{\partial^2}{\partial \theta^2} \log L(x \mid \theta)\right]_{\theta=\hat{\theta}} \xrightarrow{n \to \infty} \frac{1}{n}\left[\frac{\partial^2}{\partial \theta^2} \log L(x \mid \theta)\right]_{\theta=\theta_0}. \quad (18.23)$$

Величина

$$\frac{1}{n} \frac{\partial}{\partial \theta^2} \log L(x \mid \theta) = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta^2} \log f(x_i \mid \theta)$$

есть среднее n независимых одинаково распределенных случайных величин и согласно усиленному закону больших чисел сходится с вероятностью единица к своему математическому ожиданию. Поэтому (18.23) можно записать в виде

$$\frac{1}{n}\left[\frac{\partial^2}{\partial \theta^2} \log L(x \mid \theta)\right]_{\theta=\hat{\theta}} \xrightarrow{n \to \infty} \mathsf{M}_0\left[\frac{\partial^2}{\partial \theta^2} \log f(x \mid \theta)\right]_{\theta=\theta_0}. \quad (18.24)$$

Но согласно (17.19) при выполнении условий регулярности

$$\mathsf{M}\left[\frac{\partial^2}{\partial \theta^2} \log f(x \mid \theta)\right] = -\mathsf{M}\left\{\left(\frac{\partial \log f(x \mid \theta)}{\partial \theta}\right)^2\right\} < 0. \quad (18.25)$$

Следовательно, (18.24) дает

$$\lim_{n\to\infty} \mathbf{P}\left\{\left[\frac{\partial^2}{\partial\theta^2}\log L(x\mid\theta)\right]_{\theta=\hat{\theta}} < 0\right\} = 1. \qquad (18.26)$$

18.12 Предположим теперь, что выполнены условия пункта **18.2** и имеется два локальных максимума ФП $\hat{\theta}_1$ и $\hat{\theta}_2$, которые являются корнями (18.5) и удовлетворяют (18.6). Если $\log L(x|\theta)$ всюду имеет вторую производную, как мы предполагали в предыдущем пункте, то между двумя максимумами $\hat{\theta}_1$ и $\hat{\theta}_2$ должен находиться один минимум (обозначим его $\hat{\theta}_3$), для которого

$$\left[\frac{\partial^2 \log L(x\mid\theta)}{\partial\theta^2}\right]_{\theta=\theta_3} \geqslant 0. \qquad (18.27)$$

Но поскольку $\hat{\theta}_1$ и $\hat{\theta}_2$ суть состоятельные оценки, то оценка $\hat{\theta}_3$, значение которой заключено между ними, должна быть также состоятельной и удовлетворять (18.26). Однако (18.26) и (18.27) противоречат друг другу. Следовательно, мы можем получить только одну состоятельную оценку, являющуюся корнем уравнения правдоподобия (18.5).

18.13 В связи с рассмотрением состоятельности МП-оценок следовало бы обратить внимание на следующий факт. Не исключена возможность, что ФП имеет два (или более) максимума, т. е. в (18.2) имеет место знак равенства. Какое из значений $\hat{\theta}_1$, $\hat{\theta}_2$ и т. д., соответствующих этим максимумам, мы должны выбрать? Появляющаяся неопределенность кажется довольно серьезной. К счастью, она не играет важной роли, так как обычно трудности возникают только для отдельных комбинаций выборочных значений, вероятность появления которых мала. Однако если параметр неопределяем по существу, то неоднозначность может возникнуть для всех выборок. Следующий пример иллюстрирует такой случай.

Пример 18.4

Положим в примере 18.3

$$\cos\theta = \rho.$$

Тогда каждому действительному решению кубического уравнения правдоподобия $\hat{\rho}$ будет соответствовать бесконечное число оценок θ вида

$$\hat{\theta}_r = \arccos\hat{\rho} + 2r\pi,$$

где r — любое целое число. Параметр θ здесь в принципе невозможно оценить. Рассматриваемая как функция от θ, ФП является периодической с бесконечным числом максимумов в точках $\hat{\theta}_r$, разности между которыми кратны 2π. Конечно, для

истинного значения θ_0 может существовать только одна состоятельная оценка, однако у нас нет никакого способа определить, какая именно из оценок $\hat{\theta}_r$ состоятельна. В таких случаях мы говорим, что сам параметр *неидентифицируем* и что можно оценить лишь $\cos\theta$.

Состоятельность и смещение МП-оценок

18.14 Хотя при выполнении условий пункта **18.10** МП-оценки состоятельны, они в общем случае могут быть смещенными. Пример 18.1 показывает, что МП-оценка может быть смещенной даже тогда, когда она является функцией от одномерной достаточной статистики. Вообще говоря, этого следовало ожидать. В **18.9** мы видели, что если $\hat{\theta}$ есть МП-оценка для θ, то МП-оценкой для $\tau(\theta)$ будет $\tau(\hat{\theta})$. Однако в общем случае

$$\mathbf{M}\{\tau(\hat{\theta})\} \neq \tau\{\mathbf{M}(\hat{\theta})\}, \tag{18.28}$$

так что если $\hat{\theta}$ есть несмещенная оценка для θ, то $\tau(\hat{\theta})$ не может быть несмещенной оценкой для $\tau(\theta)$. Если МП-оценка состоятельна, то для нее можно повторить рассуждения, следующие за примером 17.3.

Бриллинджер (1964) показал, что если смещение МП-оценки уменьшено с помощью метода пункта **17.10**, то t'_n в (17.10) при выполнении условий регулярности асимптотически нормальна. Он дает также разложения для ее смещения и среднего квадрата ошибки.

Эффективность и асимптотическая нормальность МП-оценок

18.15 Рассматривая эффективность МП-оценок, мы уже не получаем таких же четких результатов, как в **18.10**. Следующий пример с достаточной ясностью показывает, что для получения оптимальных свойств МП-оценок, связанных с эффективностью, следует ввести некоторые ограничения.

Пример 18.5

В примере 17.22 было показано, что в случае распределения

$$dF(x) = dx/\theta, \qquad k\theta \leqslant x \leqslant (k+1)\theta, \qquad k > 0,$$

для θ не имеется единственной достаточной статистики, но пара крайних наблюдений $x_{(1)}$ и $x_{(n)}$ совместно достаточна для θ. Найдем теперь МП-оценку для θ. Для этого надо максимизировать ФП примера 18.1. Так как

$$L(x\,|\,\theta) = \theta^{-n} u(x_{(1)} - k\theta)\, u((k+1)\theta - x_{(n)}),$$

то

$$\hat{\theta} = x_{(n)}/(k+1)$$

будет МП-оценкой. Мы видим, что $\hat{\theta}$ зависит только от $x_{(n)}$, хотя для достаточности необходимы обе статистики $x_{(1)}$ и $x_{(n)}$.

Далее, из соображений симметрии следует, что $x_{(1)}$ и $x_{(n)}$ имеют одну и ту же дисперсию, скажем V. Дисперсия МП-оценки равна

$$\mathsf{D}\hat{\theta} = V/(k+1)^2,$$

а для дисперсии оценки

$$\theta^* = x_{(1)}/k$$

имеем

$$\mathsf{D}\theta^* = V/k^2.$$

Так как $x_{(1)}$ и $x_{(n)}$ асимптотически независимы (см. **14.23**), то функция

$$\bar{\theta} = a\hat{\theta} + (1-a)\,\theta^*$$

будет, подобно $\hat{\theta}$ и θ^*, состоятельной оценкой для θ, причем ее дисперсия

$$\mathsf{D}\bar{\theta} = V\left\{\frac{a^2}{(k+1)^2} + \frac{(1-a)^2}{k^2}\right\}$$

обращается в минимум (см. упражнение 17.21) при $a = \frac{(k+1)^2}{k^2+(k+1)^2}$. Этот минимум равен

$$\mathsf{D}\bar{\theta} = V/\{k^2 + (k+1)^2\}.$$

Таким образом, для всех $k > 0$

$$\frac{\mathsf{D}\bar{\theta}}{\mathsf{D}\hat{\theta}} = \frac{(k+1)^2}{k^2+(k+1)^2} < 1,$$

т. е. МП-оценка имеет бо́льшую дисперсию. При больших k дисперсия $\hat{\theta}$ примерно в два раза больше дисперсии $\bar{\theta}$.

18.16 Теперь мы покажем, следуя Крамеру (1946), что если ФП имеет две производные по θ в интервале, включающем истинное значение θ_0, если, кроме того,

$$\mathsf{M}\left(\frac{\partial \log L(x\,|\,\theta)}{\partial\theta}\right) = 0 \qquad (18.29)$$

и математическое ожидание

$$R^2(\theta) = -\,\mathsf{M}\left(\frac{\partial^2 \log L(x\,|\,\theta)}{\partial\theta^2}\right) = \mathsf{M}\left\{\left(\frac{\partial \log L}{\partial\theta}\right)^2\right\} \qquad (18.30)$$

существует и отлично от нуля в указанном интервале, то МП-оценка $\hat{\theta}$ асимптотически нормально распределена со средним θ_0 и дисперсией $1/R^2(\theta)$.

По формуле Тейлора имеем

$$\left(\frac{\partial \log L}{\partial \theta}\right)_{\hat{\theta}} = \left(\frac{\partial \log L}{\partial \theta}\right)_{\theta_0} + (\hat{\theta} - \theta_0)\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)_{\theta^*}, \qquad (18.31)$$

где θ^* заключено между $\hat{\theta}$ и θ_0. При наших предположениях регулярности $\hat{\theta}$ является корнем (18.5), поэтому левая часть (18.31) равна нулю, и мы можем написать

$$(\hat{\theta} - \theta_0)\, R(\theta_0) = \frac{\left(\frac{\partial \log L}{\partial \theta}\right)_{\theta_0} \Big/ R(\theta_0)}{\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)_{\theta^*} \Big/ \{-R^2(\theta_0)\}}. \qquad (18.32)$$

Так как $\hat{\theta}$ — состоятельная оценка для θ_0, а θ^* находится между $\hat{\theta}$ и θ_0, то из (18.24) и (18.30) следует, что знаменатель в правой части (18.32) стремится к единице:

$$\left[\frac{\partial^2 \log L}{\partial \theta^2}\right]_{\theta^*} \Big/ \{-R^2(\theta_0)\} \to 1. \qquad (18.33)$$

Числитель же есть поделенная на $R(\theta_0)$ сумма n независимых одинаково распределенных случайных величин $\partial \log f(x_i|\theta_0)/\partial\theta$. Эта сумма имеет среднее нуль (в силу (18.29)) и дисперсию $R^2(\theta_0)$, даваемую (18.30). Следовательно, согласно центральной предельной теореме (**7.26**) числитель будет асимптотически нормально распределен со средним 0 и дисперсией 1. То же самое можно сказать о всей правой части (18.32). Таким образом, левая часть (18.32) асимптотически является стандартной нормальной величиной. Другими словами, МП-оценка $\hat{\theta}$ асимптотически нормально распределена со средним θ_0 и дисперсией $1/R^2(\theta_0)$.

Даниэлс (1961) ослабил сформулированное выше условие асимптотической нормальности и эффективности МП-оценок.

18.17 Из приведенного результата, дающего для МП-оценки асимптотическую дисперсию, равную МГД, вытекает, что при сформулированных выше условиях МП-оценка эффективна. Поскольку МГД может быть достигнута только при наличии достаточной статистики (см. **17.33**), то правомерно сказать, что МП-оценка «асимпотически достаточна».

Ле Кам (1953) возражал против использования термина «эффективный», поскольку с ним связано представление о достижении абсолютного минимума дисперсии для больших выборок, а этот минимум, строго говоря, не достигается ни для МП-оценок, ни для каких-либо других оценок. Рассмотрим, например, состоятельную, асимптотически нормально распределенную с дисперсией порядка n^{-1} оценку t для θ. Определим новую статистику

$$t' = \begin{cases} t, & \text{если } |t| \geqslant n^{-1/4}, \\ kt, & \text{если } |t| < n^{-1/4}. \end{cases} \qquad (18.34)$$

Поскольку

$$\lim_{n\to\infty} \mathrm{D}t'/\mathrm{D}t = \begin{cases} 1, & \text{если} \quad \theta \neq 0, \\ k^2, & \text{если} \quad \theta = 0, \end{cases}$$

то при $k < 1$ оценка t' будет для $\theta = 0$ более эффективной, чем t, а для других значений θ будет столь же эффективной. Ле Кам показал (см. также Бахадур (1964)), что подобная «сверхэффективность» может иметь место только для значений θ, образующих множество меры нуль. Учитывая это, мы сохраним термин «эффективный» в его обычном смысле. С. Р. Рао (1962b), однако, показывает, что даже этого не очень существенного парадокса можно избежать, если определить эффективность оценки в терминах ее корреляции с $\dfrac{\partial \log L}{\partial \theta}$ (ср. **17.15—17** и (17.61)). Уокер (1963) дает условия регулярности, достаточные для того, чтобы МГД ограничивала асимптотические дисперсии всех асимптотически нормальных оценок.

Пример 18.6

В примере 18.3 было показано, что МП-оценка $\hat{\rho}$ для параметра корреляции нормированного двумерного нормального распределения является корнем кубического уравнения

$$\frac{\partial \log L}{\partial \rho} = \frac{n\rho}{1-\rho^2} - \frac{\rho}{(1-\rho^2)^2}\left\{\sum x^2 - 2\rho \sum xy + \sum y^2\right\} + \\ + \frac{1}{1-\rho^2}\sum xy = 0.$$

Дифференцируя второй раз, получим

$$\frac{\partial^2 \log L}{\partial \rho^2} = \frac{n(1+\rho^2)}{(1-\rho^2)^2} - \frac{(1+3\rho^2)}{(1-\rho^2)^3}\left(\sum x^2 - 2\rho \sum xy + \sum y^2\right) + \\ + \frac{4\rho}{(1-\rho^2)^2}\sum xy$$

или, учитывая, что $\mathbf{M}(x^2) = \mathbf{M}(y^2) = 1$ и $\mathbf{M}(xy) = \rho$,

$$\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \rho^2}\right) = \frac{n(1+\rho^2)}{(1-\rho^2)^2} - \frac{2n(1+3\rho^2)}{(1-\rho^2)^2} + \frac{4n\rho^2}{(1-\rho^2)^2} = -\frac{n(1+\rho^2)}{(1-\rho^2)^2}.$$

Следовательно, согласно **18.16** асимптотически имеем

$$\mathrm{D}\hat{\rho} = \frac{(1-\rho^2)^2}{n(1+\rho^2)}.$$

Пример 18.7

Распределению

$$dF(x) = \frac{1}{2}\exp\{-|x-\theta|\}\,dx, \qquad -\infty \leqslant x \leqslant \infty,$$

соответствует логарифм правдоподобия

$$\log L(x\,|\,\theta) = -n\log 2 - \sum_{i=1}^{n} |x_i - \theta|,$$

который максимален, когда сумма $\sum_i |x_i - \theta|$ минимальна, т. е. когда θ является медианой x_i (см. упражнение 2.1). (Если n нечетно, то медианой будет среднее наблюдение; если же n четно, то в качестве медианы можно взять любое значение из интервала, концами которого служат два средних наблюдения.) Таким образом, МП-оценкой является выборочная медиана, $\hat{\theta} = \tilde{x}$. Из (14.20) легко видеть, что ее дисперсия для больших выборок равна $\mathsf{D}\hat{\theta} = 1/n$.

В данном случае нельзя использовать результаты пункта **18.16** для проверки эффективности $\hat{\theta}$, поскольку требуемые условия дифференцируемости для рассматриваемого распределения не выполняются. Но поскольку производная

$$\frac{\partial \log f(x \mid \theta)}{\partial \theta} = \begin{cases} +1, & \text{если } x > \theta, \\ -1, & \text{если } x < \theta, \end{cases}$$

не существует лишь при $x = \theta$, мы имеем

$$\left(\frac{\partial \log f(x \mid \theta)}{\partial \theta}\right)^2 = 1, \qquad x \neq \theta,$$

так что если интерпретировать $\mathsf{M}\left[\left(\frac{\partial \log f(x \mid \theta)}{\partial \theta}\right)^2\right]$ как

$$\lim_{\varepsilon \to 0} \left\{ \int_{-\infty}^{\theta-\varepsilon} + \int_{\theta+\varepsilon}^{\infty} \right\} 1 \cdot dF(x) = 1,$$

то получим

$$\mathsf{M}\left\{\left(\frac{\partial \log L}{\partial \theta}\right)^2\right\} = n\mathsf{M}\left\{\left(\frac{\partial \log f(x \mid \theta)}{\partial \theta}\right)^2\right\} = n.$$

Отсюда следует, что МГД оценок для θ равна $\mathsf{D}t \geqslant 1/n$. Как было показано, эта граница асимптотически достигается для $\hat{\theta}$.

18.18 Результат пункта **18.16** упрощается в случае распределения, допускающего для параметра одномерную достаточную статистику. Поскольку тогда из (18.10), (18.11) и (18.12) получаем

$$\mathsf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right) = -A(\theta)\,\tau'(\theta) = \left(\frac{\partial^2 \log L}{\partial \theta^2}\right)_{\theta=\theta}, \qquad (18.35)$$

то отпадает необходимость в вычислении математического ожидания. МГД в этом случае равна просто $-1\Big/\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)_{\theta=\theta}$ и достигается точно, если $\hat{\theta}$ является несмещенной оценкой для θ, и асимптотически в случае выполнения условий пункта **18.16**. Если не существует одномерной достаточной статистики, то асимптотическая дисперсия $\hat{\theta}$ может быть оценена по выборке обычным образом, при этом как правило ищут несмещенную оценку.

Пример 18.8

Оценить стандартное отклонение σ нормального распределения

$$dF(x) = \frac{1}{\sigma(2\pi)^{1/2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx, \qquad -\infty \leqslant x \leqslant \infty.$$

Поскольку

$$\log L(x \mid \theta) = -n \log \sigma - \frac{\sum x^2}{2\sigma^2}, \quad (\log L)' = -n/\sigma + \sum x^2/\sigma^2,$$

то достаточной МП-оценкой будет $\hat{\sigma} = \sqrt{\frac{1}{n}\sum x^2}$ и

$$(\log L)'' = \frac{n}{\sigma^2} - \frac{3}{\sigma^4}\sum x^2 = \frac{n}{\sigma^2}\left(1 - \frac{3\hat{\sigma}^2}{\sigma^2}\right).$$

Таким образом, используя (18.35), мы получим при возрастании n

$$\mathrm{D}\hat{\sigma} \sim -1/(\log L)''_{\hat{\sigma}=\sigma} = \sigma^2/(2n).$$

Семиинварианты одной МП-оценки

18.19 Холдейн и Смит (1956) получили, при некоторых условиях регулярности, выражения для первых четырех семиинвариантов МП-оценки в следующей ситуации. Предположим, что область определения распределения, из которого производится выбор, разбита на счетное число классов и вероятность попадания наблюдения в r-й класс равна π_r ($r = 1, 2, \ldots$). Мы сводим, таким образом, любое распределение к мультиномиальному (**5.30**), и если область определения исходного распределения не зависит от параметра θ, то мы можем искать решения уравнения правдоподобия (18.5). Поскольку вероятности π_r являются функциями от θ, то из (5.78) получим

$$L(x \mid \theta) \propto \prod_r \pi_r^{n_r}, \tag{18.36}$$

где n_r — число наблюдений в r-м классе и $\sum_r n_r = n$ — объем выборки. Согласно (18.36) уравнение правдоподобия можно записать в виде

$$\frac{\partial \log L}{\partial \theta} = \sum_r n_r \frac{\pi_r'}{\pi_r} = 0, \tag{18.37}$$

где штрих означает дифференцирование по θ. Разлагая π_r и π_r' по формуле Тейлора в окрестности истинного значения θ_0, получим

$$\left.\begin{aligned}\pi_r(\hat{\theta}) &= \pi_r(\theta_0) + (\hat{\theta}-\theta_0)\,\pi_r'(\theta_0) + \tfrac{1}{2}(\hat{\theta}-\theta_0)^2\pi_r''(\theta_0) + \ldots,\\ \pi_r'(\hat{\theta}) &= \pi_r'(\theta_0) + (\hat{\theta}-\theta_0)\,\pi_r''(\theta_0) + \tfrac{1}{2}(\hat{\theta}-\theta_0)^2\pi_r'''(\theta_0) + \ldots\end{aligned}\right\} \quad (18.38)$$

Подстановка (18.38) в (18.37), биномиальное разложение и суммирование полученных рядов приводят к разложению

$$\alpha_1 - (A_1 + \alpha_2 - \beta_1)(\hat{\theta}-\theta_0) + \tfrac{1}{2}(2A_2 - 3B_1 + 2\alpha_3 - 3\beta_2 + \delta_1)(\hat{\theta}-\theta_0)^2 - \\ - \tfrac{1}{6}(6A_3 - 12B_2 + 3C_1 + 4D_1)(\hat{\theta}-\theta_0)^3 + \ldots = 0. \quad (18.39)$$

В (18.39) использованы следующие обозначения:

$$A_i = \sum_r \{\pi_r'(\theta_0)\}^{i+1}/\{\pi_r(\theta_0)\}^i,$$

$$B_i = \sum_r \{\pi_r'(\theta_0)\}^i\,\pi_r''(\theta_0)/\{\pi_r(\theta_0)\}^i,$$

$$C_i = \sum_r \{\pi_r'(\theta_0)\}^{i-1}\{\pi_r''(\theta_0)\}^2/\{\pi_r(\theta_0)\}^i,$$

$$D_i = \sum_r \{\pi_r'(\theta_0)\}^i\,\pi_r'''(\theta_0)/\{\pi_r(\theta_0)\}^i,$$

$$\alpha_i = \sum_r \{\pi_r'(\theta_0)\}^i\left\{\frac{n_r}{n} - \pi_r(\theta_0)\right\}/\{\pi_r(\theta_0)\}^i,$$

$$\beta_i = \sum_r \{\pi_r'(\theta_0)\}^{i-1}\,\pi_r''(\theta_0)\left\{\frac{n_r}{n} - \pi_r(\theta_0)\right\}/\{\pi_r(\theta_0)\}^i,$$

$$\delta_i = \sum_r \{\pi_r'(\theta_0)\}^{i-1}\,\pi_r'''(\theta_0)\left\{\frac{n_r}{n} - \pi_r(\theta_0)\right\}/\{\pi_r(\theta_0)\}^i.$$

При больших n можно применить теорему Лагранжа и обратить (18.39), что дает

$$(\hat{\theta}-\theta_0) = A_1^{-1}\alpha_1 + A_1^{-3}\alpha_1\left[\left(A_2 - \tfrac{3}{2}B_1\right)\alpha_1 - A_1(\alpha_2-\beta_1)\right] + \\ + A_1^{-5}\alpha_1\left[\left\{2\left(A_2 - \tfrac{3}{2}B_1\right)^2 - A_1\left(A_3 - 2B_2 + \tfrac{1}{2}C_1 + \tfrac{2}{3}D_1\right)\right\}\alpha_1^2 - \right. \\ - 3A_1\left(A_2 - \tfrac{3}{2}B_1\right)\alpha_1(\alpha_2-\beta_1) + \\ \left. + \tfrac{1}{2}A_1^2\alpha_1(2\alpha_3 - 3\beta_2 + \delta_1) + A_1^2(\alpha_2-\beta_1)^2\right] + O(n^{-3}). \quad (18.40)$$

(18.40) позволяет нам получить моменты $\hat{\theta}$ в виде рядов по степеням n^{-1}.

18.20 Рассмотрим выборочное распределение суммы

$$W = \sum_r h_r \left\{ \frac{n_r}{n} - \pi_r(\theta_0) \right\},$$

где h_r — произвольные постоянные веса. Используя формулы для моментов мультиномиального распределения (см. (5.80)) и обозначая $S_i = \sum_r h_r^i \pi_r(\theta_0)$, получаем моменты W:

$$\left.\begin{aligned}
&\mu_1'(W) = 0,\\
&\mu_2(W) = n^{-1}(S_2 - S_1^2),\\
&\mu_3(W) = n^{-2}(S_3 - S_1S_2 + 2S_1^3),\\
&\mu_4(W) = 3n^{-2}(S_2 - S_1^2)^2 +\\
&\qquad + n^{-3}(S_4 - 4S_1S_3 - 3S_2^2 + 12S_1^2S_2 - 6S_1^4),\\
&\mu_5(W) = 10n^{-3}(S_2 - S_1^2)(S_3 - 3S_1S_2 + 2S_1^3) + O(n^{-4}),\\
&\mu_6(W) = 15n^{-3}(S_2 - S_1^2)^3 + O(n^{-4}).
\end{aligned}\right\} \qquad (18.41)$$

Отсюда можно получить выражения для моментов и смешанных моментов случайных величин α_i, β_i и δ_i, фигурирующих в (18.40), поскольку все они являются функциями вида W. Наконец, подставляя эти моменты в (18.40), мы получим моменты $\hat{\theta}$. Выражая через них семиинварианты, находим

$$\left.\begin{aligned}
&\varkappa_1 = \theta_0 - \tfrac{1}{2} n^{-1} A_1^{-2} B_1 + O(n^{-2}),\\
&\varkappa_2 = n^{-1}A_1^{-1} + n^{-2}A_1^{-4}\left[-A_2^2 + \tfrac{7}{2}B_1^2 + \right.\\
&\qquad \left. + A_1(A_3 - B_2 - D_1) - A_1^3 \right] + O(n^{-3}),\\
&\varkappa_3 = n^{-2}A_1^{-3}(A_2 - 3B_1) + O(n^{-3}),\\
&\varkappa_4 = n^{-3}A_1^{-5}\left[-12B_1(A_2 - 2B_1) + A_1(A_3 - 4D_1) - 3A_1^3 \right] +\\
&\qquad + O(n^{-4}),
\end{aligned}\right\} \qquad (18.42)$$

откуда

$$\left.\begin{aligned}
&\gamma_1 = \varkappa_3/\varkappa_2^{3/2} = n^{-1/2}A_1^{-3/2}(A_2 - 3B_1) + o(n^{-1/2}),\\
&\gamma_2 = \varkappa_4/\varkappa_2^2 = n^{-1}A_1^{-3}\left[-12B_1(A_2 - 2B_1) + \right.\\
&\qquad \left. + A_1(A_3 - 4D_1) - 3A_1^3 \right] + o(n^{-1}).
\end{aligned}\right\} \qquad (18.43)$$

Первый семиинвариант в (18.42) показывает, что $\hat{\theta}$ имеет смещение порядка n^{-1}, если $B_1 \neq 0$. Если $B_1 = 0$, то смещение имеет порядок n^{-2}, что можно показать, вычисляя следующий член в разложении первого семиинварианта. Главный член во втором семиинварианте есть просто асимптотическая дисперсия, полученная в **18.16**. (18.43) показывает скорость установленной в **18.16** сходимости к нормальному закону.

Если вычислить все члены в (18.42) и получить затем несмещенные оценки первых четырех моментов $\hat{\theta}$, то можно с помощью распределения Пирсона (см. **6.2—12**) приблизить распределение $\hat{\theta}$ при малых объемах выборок. Это приближение было бы лучше, чем нормальное приближение, полученное в **18.16**.

Шентоном и Боумэном (1963) были получены члены более высокого порядка в (18.42—3).

Последовательное приближение к МП-оценке

18.21 В большинстве рассмотренных нами примеров МП-оценки были получены в явном виде. Исключением является пример 18.3, где мы пришли к кубическому уравнению. Но даже в этом случае не представляет особого труда получить решение, когда значения x заданы. Однако иногда уравнение правдоподобия настолько сложно, что приходится искать корень с помощью итерационных методов, начиная с некоторого пробного значения t.

Как и в (18.31), мы разлагаем $\partial \log L/\partial\theta$ в ряд Тейлора (но на этот раз в окрестности t) и получаем

$$0 = \left(\frac{\partial \log L}{\partial \theta}\right)_{\hat{\theta}} = \left(\frac{\partial \log L}{\partial \theta}\right)_t + (\hat{\theta} - t)\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)_{\theta^*}$$

(θ^* лежит между $\hat{\theta}$ и t). Отсюда

$$\hat{\theta} = t - \left(\frac{\partial \log L}{\partial \theta}\right)_t \Big/ \left(\frac{\partial^2 \log L}{\partial \theta^2}\right)_{\theta^*}. \tag{18.44}$$

Если некоторое t не очень сильно отличается от $\hat{\theta}$, то можно заменить θ^* в (18.44) на t. Это дает

$$\hat{\theta} = t - \left(\frac{\partial \log L}{\partial \theta}\right)_t \Big/ \left(\frac{\partial^2 \log L}{\partial \theta^2}\right)_t, \tag{18.45}$$

что является лучшим приближением для $\hat{\theta}$, чем t. Этот процесс можно повторять до тех пор, пока не будет получена желаемая точность.

Обычно в качестве t берут значение какой-нибудь легко вычисляемой состоятельной оценки θ. Тогда при $n \to \infty$ мы бу-

дем иметь две состоятельные оценки t и $\hat{\theta}$, сходящиеся к θ_0. Очевидно, θ^* также будет сходиться к θ_0. Три случайные величины $\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)_{\theta^*}$, $\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)_t$ и $\left[\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)\right]_t$ будут сходиться к $\left[\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)\right]_{\theta_0}$. Использование в (18.44) второй из этих величин вместо первой дает (18.45). Использование же третьей вместо первой дает другую итерационную процедуру:

$$\hat{\theta} = t - \left(\frac{\partial \log L}{\partial \theta}\right)_t \Big/ \left[\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)\right]_t = t + \left(\frac{\partial \log L}{\partial \theta}\right)_t (\mathbf{D}\hat{\theta})_t \quad (18.46)$$

(последнее равенство следует из **18.16**). (18.45) есть итерационный процесс Ньютона — Рафсона, а процедура (18.46) предложена Фишером (1925). Кейл (1961) показал, что (18.46) при больших n обычно быстрее приводит к цели, если не требуется чрезмерно высокая окончательная точность. Обычно процедура (18.46) менее трудоемка.

Как (18.45), так и (18.46) могут в некоторых случаях не сходиться. Если даже они сходятся, но уравнение правдоподобия имеет несколько корней, то нет никакой гарантии, что они сходятся к корню, соответствующему абсолютному максимуму ФП. Поэтому следует найти интервалы, в которых $\partial \log L/\partial\theta$ меняет знак с плюса на минус, вычислить все максимумы и затем сравнить их. Барнетт (1966а) рассматривает систематический метод выполнения такой операции, используя «метод ложных положений».

Пример 18.9

Оценить параметр θ в распределении Коши

$$dF(x) = \frac{dx}{\pi\{1 + (x-\theta)^2\}}, \qquad -\infty \leqslant x \leqslant \infty.$$

Уравнение правдоподобия

$$\frac{\partial \log L}{\partial \theta} = 2\sum_{i=1}^{n} \frac{(x_i - \theta)}{\{1 + (x_i - \theta)^2\}} = 0$$

есть уравнение степени $(2n-1)$ по θ. Согласно **18.16** для асимптотической дисперсии $\hat{\theta}$ имеем

$$-\frac{1}{\mathbf{D}\hat{\theta}} \sim \mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right) = n\mathbf{M}\left(\frac{\partial^2 \log f}{\partial \theta^2}\right) =$$

$$= \frac{n}{\pi}\int_{-\infty}^{\infty} \frac{2(x-\theta)^2 - 2}{\{1 + (x-\theta)^2\}^3}\, dx = \frac{4n}{\pi}\int_0^{\infty} \frac{(x^2-1)\,dx}{(1+x^2)^3} = -\frac{n}{2}.$$

Следовательно,

$$\mathrm{D}\hat{\theta} = 2/n.$$

Уравнение правдоподобия имеет в общем случае несколько корней, и при малых n (18.45) или (18.46) могут не сходиться (см. Барнетт (1966a)). Однако с вероятностью, стремящейся к 1, $\hat{\theta}$ является максимумом, ближайшим к выборочной медиане, асимптотическая дисперсия которой равна (пример 17.5)

$$\mathrm{D}t = \pi^2/(4n),$$

а эффективность $8/\pi^2 \approx 0{,}8$. Следовательно, при больших n мы можем взять в качестве начальной точки при отыскании $\hat{\theta}$ выборочную медиану и использовать формулу (18.46), которая в данном случае примет вид

$$\hat{\theta} = t + \frac{4}{n}\sum_i \frac{(x_i - t)}{\{1 + (x_i - t)^2\}}.$$

Это будет первым приближением к $\hat{\theta}$. Более точные приближения можно получить с помощью дальнейших итераций.

Блоч (1966) предложил в качестве оценки для θ линейную комбинацию пяти порядковых статистик $x_{(r)}$ ($r = 0{,}13n$; $0{,}4n$; $0{,}5n$; $0{,}6n$; $0{,}87n$) с весами —0,052; 0,3485; 0,407; 0,3485; —0,052. Эффективность этой оценки превышает 0,95, поэтому она была бы превосходной начальной точкой при итерационном решении уравнения правдоподобия. Ротенберг и др. (1964) показали, что среднее значение центральных 24 процентов выборки из распределения Коши имеет асимптотическую дисперсию $2{,}28/n$ и, следовательно, эффективность 0,88. См. также Барнетт (1966b).

Пример 18.10

Мы рассмотрим теперь итерационный метод решения более детально и для этой цели воспользуемся некоторыми данными, полученными Фишером (1925, глава 9).

Рассмотрим мультиномиальное распределение с четырьмя классами, имеющими вероятности

$$p_1 = (2 + \theta)/4, \quad p_2 = p_3 = (1 - \theta)/4, \quad p_4 = \theta/4.$$

Параметр θ, лежащий в интервале $(0, 1)$, должен быть оценен исходя из наблюденных частот (a, b, c, d), соответствующих четырем классам (объем выборки n равен $a + b + c + d$). Имеем

$$L(a, b, c, d \mid \theta) \propto (2 + \theta)^a (1 - \theta)^{b+c} \theta^d,$$

откуда

$$\frac{\partial \log L}{\partial \theta} = \frac{a}{2 + \theta} - \frac{b + c}{1 - \theta} + \frac{d}{\theta}.$$

Приравнивая последнее выражение нулю, получаем квадратное уравнение относительно θ:

$$n\theta^2 + \{2(b+c) + d - a\}\theta - 2d = 0.$$

Поскольку произведение коэффициента при θ^2 на свободный член отрицательно, то произведение корней квадратного уравнения должно быть также отрицательно; поэтому только один корень может быть положительным, т. е. лежать в области возможных значений θ. Значение этого корня $\hat{\theta}$ дается выражением

$$2n\hat{\theta} = \{a - d - 2(b+c)\} + [\{a + 2(b+c) + 3d\}^2 - 8a(b+c)]^{1/2}.$$

Полученная формула позволяет очень просто вычислять МП-оценку $\hat{\theta}$. В примере Фишера (из генетики) наблюденные частоты были равны

$$a = 1997, \quad b = 906, \quad c = 904, \quad d = 32, \quad n = 3839,$$

а для $\hat{\theta}$ было получено значение 0,0357.

Легко показать, проводя еще одно дифференцирование, что

$$\mathrm{D}\hat{\theta} \sim -\frac{1}{\mathrm{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)} = \frac{2\theta(1-\theta)(2+\theta)}{n(1+2\theta)}.$$

Подставляя $\hat{\theta}$ вместо θ, мы получаем в рассматриваемом случае для $\mathrm{D}\hat{\theta}$ значение 0,0000336.

В целях иллюстрации мы найдем $\hat{\theta}$ для данного примера итерационным путем, начав с какой-нибудь неэффективной оценки. Фишером была предложена следующая простая неэффективная оценка:

$$t = \{a + d - (b+c)\}/n.$$

Легко видеть, что эта оценка состоятельна и имеет дисперсию

$$\mathrm{D}t = (1 - \theta^2)/n.$$

Значение t для генетических данных равно

$$t = \{1997 + 32 - (906 + 904)\}/3839 = 0{,}0570.$$

Это довольно далеко от значения $\hat{\theta}$, равного 0,0357, которое мы ищем. Используя (18.46), получаем первое приближение для $\hat{\theta}$:

$$\hat{\theta}_1 = 0{,}0570 + \left(\frac{\partial \log L}{\partial \theta}\right)_{\theta=t} (\mathrm{D}\hat{\theta})_{\theta=t}.$$

Поскольку

$$\left(\frac{\partial \log L}{\partial \theta}\right)_{\theta=0,0570} = \frac{1997}{2,0570} - \frac{1810}{0,9430} + \frac{32}{0,0570} = -387,1713,$$

$$(D\hat{\theta})_{\theta=0,0570} = \frac{2 \times 0,0570 \times 0,943 \times 2,057}{3839 \times 1,114} = 0,00005170678,$$

то наша улучшенная оценка равна

$$\hat{\theta}_1 = 0,0570 - 387,1713 \times 0,00005170678 = 0,0570 - 0,0200 = 0,0370,$$

т. е. довольно близка к искомой величине 0,0357. Вторая итерация дает

$$\left(\frac{\partial \log L}{\partial \theta}\right)_{\theta=0,0370} = \frac{1997}{2,037} - \frac{1810}{0,963} + \frac{32}{0,037} = -34,31495,$$

$$(D\hat{\theta})_{\theta=0,0370} = \frac{2 \times 0,037 \times 0,963 \times 2,037}{3839 \times 1,074} = 0,00003520681.$$

Следовательно,

$$\hat{\theta}_2 = 0,0370 - 34,31495 \times 0,00003520681 = 0,0370 - 0,0012 = 0,0358.$$

Полученное значение очень близко к искомому. Чтобы получить значение 0,0357, верное с точностью до четырех десятичных знаков, нужна еще одна итерация. Кроме того, еще одна итерация необходима, чтобы подтвердить, что полученное значение $\hat{\theta}$ устойчиво в пределах требуемого числа десятичных знаков, т. е. дальнейшие итерации не нужны. Читателю следовало бы сделать эти итерации, чтобы убедиться в своем умении использовать описанный метод.

Рассмотренный пример показывает, что при использовании итерационного метода для практических целей следует позаботиться о том, чтобы было сделано достаточное число итераций. Правда, этот пример является неблагоприятным, поскольку эффективность t низка. Она равна $\frac{2\theta(2+\theta)}{(1+\theta)(1+2\theta)}$ и при подстановке $\hat{\theta} = 0,0357$ вместо θ принимает значение 0,13 или 13 процентов. Обычно стараются начинать со значения оценки с более высокой эффективностью.

МП-оценки для нескольких параметров

18.22 Теперь мы переходим к рассмотрению общего случая, когда требуется оценить одновременно несколько параметров одномерного или многомерного распределения. Если θ и x понимать как векторы, то формулировка принципа МП (**18.2**) сохранится: надо найти такую совокупность допустимых значений параметров $\theta_1, \ldots, \theta_k$, которая обращает ФП в абсолютный

максимум. В условиях регулярности пунктов **18.2—3** система уравнений

$$\frac{\partial}{\partial\theta_r}\log L(x\,|\,\theta_1,\ \ldots,\ \theta_k)=0,\qquad r=1,\ 2,\ \ldots,\ k,\qquad (18.47)$$

дает необходимые условия локального экстремума, а неотрицательная определенность матрицы

$$\left(\frac{\partial^2\log L}{\partial\theta_r\,\partial\theta_s}\right)\qquad (18.48)$$

является достаточным условием для того, чтобы этот локальный экстремум был максимумом ФП. Решение системы (18.47) дает МП-оценки $\hat{\theta}_1,\ \ldots,\ \hat{\theta}_k$.

Случай совместной достаточности

18.23 Точно так же, как и в **18.4**, в случае существования для параметров $\theta_1,\ \ldots,\ \theta_k$ системы из s совместно достаточных статистик $t_1,\ \ldots,\ t_s$ МП-оценки $\hat{\theta}_1,\ \ldots,\ \hat{\theta}_k$ будут функциями этих статистик. Как и раньше, это немедленно следует из факторизации (см. (17.84))

$$L(x\,|\,\theta_1,\ \ldots,\ \theta_k)=g(t_1,\ \ldots,\ t_s\,|\,\theta_1,\ \ldots,\ \theta_k)\,h(x)\qquad (18.49)$$

в силу того, что $h(x)$ не содержит $\theta_1,\ \ldots,\ \theta_k$.

Однако МП-оценки не обязаны быть взаимно однозначными функциями от достаточной системы статистик и, следовательно, сами могут не составлять достаточную систему. В примере 18.5 мы уже встретились со случаем, когда МП-оценка единственного параметра была функцией только от одной из двух совместно достаточных статистик (одномерной достаточной статистики в этом примере не существовало).

18.24 Единственность решения уравнения правдоподобия в случае достаточности (**18.5**) обобщается, как показал Хузурбазар (1949), на многопараметрический случай, если $s=k$. В условиях регулярности наиболее общей форме распределения, допускающего систему из k достаточных статистик (17.86), соответствует ФП, логарифм которой равен

$$\log L=\sum_{j=1}^{k}A_j(\theta)\sum_{i=1}^{n}B_j(x_i)+\sum_{i=1}^{n}C(x_i)+nD(\theta),\qquad (18.50)$$

где вместо $\theta_1,\ \ldots,\ \theta_k$ написано θ, а x может быть вектором. Уравнения правдоподобия имеют, следовательно, вид

$$\frac{\partial\log L}{\partial\theta_r}=\sum_j\frac{\partial A_j}{\partial\theta_r}\sum_i B_j(x_i)+n\frac{\partial D}{\partial\theta_r}=0,\quad r=1,\ 2,\ \ldots,\ k.\qquad (18.51)$$

Решение $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k)$ этого уравнения дает максимум, если неотрицательно определена матрица (18.48):

$$\left(\frac{\partial^2 \log L}{\partial\theta_r\,\partial\theta_s}\right)_{\hat{\theta}} = \sum_j \left(\frac{\partial^2 A_j}{\partial\theta_r\,\partial\theta_s}\right)_{\hat{\theta}} \sum_i B_j(x_i) + n\left(\frac{\partial^2 D}{\partial\theta_r\,\partial\theta_s}\right)_{\hat{\theta}}. \quad (18.52)$$

Из (17.18) имеем

$$\mathbf{M}\left(\frac{\partial \log L}{\partial\theta_r}\right) = \sum \frac{\partial A_j}{\partial\theta_r}\,\mathbf{M}\left(\sum_i B_j(x_i)\right) + n\frac{\partial D}{\partial\theta_r} = 0. \quad (18.53)$$

Кроме того,

$$\mathbf{M}\left(\frac{\partial^2 \log L}{\partial\theta_r\,\partial\theta_s}\right) = \sum_j \frac{\partial^2 A_j}{\partial\theta_r\,\partial\theta_s}\,\mathbf{M}\left(\sum_i B_j(x_i)\right) + n\frac{\partial^2 D}{\partial\theta_r\,\partial\theta_s}. \quad (18.54)$$

Очевидно, (18.53) и (18.54) имеют точно такую же структуру, как (18.51) и (18.52), с той лишь разницей, что сумма $T = \sum_i B_j(x_i)$ заменена ее математическим ожиданием, а $\hat{\theta}$ — истинным значением θ. Если мы исключим T из (18.52), используя (18.51), и заменим $\hat{\theta}$ на θ, то мы получим такой же результат, как если бы мы исключали $\mathbf{M}(T)$ из (18.54), используя (18.53). Таким образом, имеем

$$\left(\frac{\partial^2 \log L}{\partial\theta_r\,\partial\theta_s}\right)_{\hat{\theta}=\theta} = \mathbf{M}\left(\frac{\partial^2 \log L}{\partial\theta_r\,\partial\theta_s}\right), \quad (18.55)$$

что является обобщением (18.35). Кроме того, из (17.19) получаем

$$\mathbf{M}\left(\frac{\partial^2 \log L}{\partial\theta_r^2}\right) = -\,\mathbf{M}\left\{\left(\frac{\partial \log L}{\partial\theta_r}\right)^2\right\} \quad (18.56)$$

и, аналогично,

$$\mathbf{M}\left(\frac{\partial^2 \log L}{\partial\theta_r\,\partial\theta_s}\right) = -\,\mathbf{M}\left(\frac{\partial \log L}{\partial\theta_r}\cdot\frac{\partial \log L}{\partial\theta_s}\right\}. \quad (18.57)$$

Мы видим, что матрица

$$\left\{\mathbf{M}\left(\frac{\partial^2 \log L}{\partial\theta_r\,\partial\theta_s}\right)\right\} = -\left\{\mathbf{M}\left(\frac{\partial \log L}{\partial\theta_r}\cdot\frac{\partial \log L}{\partial\theta_s}\right)\right\} = -\left\{\operatorname{cov}\left(\frac{\partial \log L}{\partial\theta_r}, \frac{\partial \log L}{\partial\theta_s}\right)\right\} \quad (18.58)$$

отрицательно или неположительно определена, так как матрица в правой части (18.58) есть матрица рассеяния $\boldsymbol{D}$ вариант $\partial \log L/\partial\theta_r$. Матрица рассеяния $\boldsymbol{D}$ любых вариант x_r неотрицательно определена, поскольку

$$\mathbf{M}\left\{\sum_{r=1}^{k}(x_r - \mathbf{M}(x_r))\,u_r\right\}^2 = \boldsymbol{u}'\,\boldsymbol{D}\boldsymbol{u} \geqslant 0, \quad (18.59)$$

где $\boldsymbol{u}$ — вектор произвольных переменных. Итак, матрица рассеяния $\boldsymbol{D}$ неотрицательно определена. Если мы исключим линейные зависимости между переменными, она станет положительно определенной, а матрица в левой части (18.55) — отрицательно определенной. Следовательно, согласно (18.55) матрица

$$\left\{\left(\frac{\partial^2 \log L}{\partial\theta_r\,\partial\theta_s}\right)_{\hat{\theta}=\theta}\right\}$$

также отрицательно определена, и поэтому любое решение (18.51) соответствует максимуму. Но в условиях регулярности между любыми двумя максимумами должен находиться минимум. Поскольку нет ни одного минимума, то может иметься только один максимум *). Таким образом, в условиях регулярности из совместной достаточности следует, что уравнение правдоподобия имеет единственное решение и что это решение соответствует максимуму ФП.

Пример 18.11

В примере 17.17 мы видели, что выборочные среднее и дисперсия $\bar{x}$ и s^2 одномерного нормального распределения совместно достаточны для генерального среднего μ и генеральной дисперсии σ^2. Из **18.23** следует, что МП-оценки должны быть функциями от $\bar{x}$ и s^2. Можно непосредственно проверить, что $\bar{x}$ и s^2 сами являются МП-оценками. ФП дается выражением

$$\log L = -\tfrac{1}{2} n \log(2\pi) - \tfrac{1}{2} n \log(\sigma^2) - \sum_i (x_i - \mu)^2/(2\sigma^2).$$

Следовательно, уравнения правдоподобия имеют вид

$$\frac{\partial \log L}{\partial \mu} = \frac{\sum (x-\mu)}{\sigma^2} = \frac{n(\bar{x}-\mu)}{\sigma^2} = 0,$$

$$\frac{\partial \log L}{\partial (\sigma^2)} = -\frac{n}{2\sigma^2} + \frac{\sum (x-\mu)^2}{2\sigma^4} = 0.$$

Их решениями будут

$$\hat{\mu} = \bar{x}, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum (x-\bar{x})^2 = s^2.$$

В то время как $\hat{\mu}$ — несмещенная оценка, $\hat{\sigma}^2$ смещена и имеет математическое ожидание $(n-1)\sigma^2/n$. Как и в однопараметрическом случае **(18.14)**, МП-оценки не обязаны быть несмещенными.

*) Утверждение о наличии минимума между двумя максимумами справедливо только при $k = 1$. Общий случай можно свести к одномерному, рассматривая семейство (18.50) с θ, пробегающим прямую, проходящую через два предполагаемых максимума. (*Прим. ред.*)

18.25 Насколько нам известно, не имеется какого-либо общего исследования единственности МП-оценок при наличии достаточных статистик для случая, когда пределы области определения распределения зависят от нескольких параметров. Однако если статистики как совместно, так и индивидуально достаточны для параметров, от которых зависят эти пределы, то результаты, полученные в **18.6** для однопараметрического случая, очевидным образом, сохраняются. Это иллюстрируется следующим примером.

Пример 18.12

В примере 17.21 было показано, что для распределения

$$dF(x) = \frac{dx}{\beta - \alpha}, \qquad \alpha \leqslant x \leqslant \beta,$$

крайние наблюдения $x_{(1)}$ и $x_{(n)}$ совместно достаточны для α и β. Ясно, что в этом случае МП-оценки

$$\hat{\alpha} = x_{(1)}, \qquad \hat{\beta} = x_{(n)}$$

являются единственными, которые максимизируют ФП. Такой результат будет всегда, когда границы интервала, на котором сосредоточено распределение, зависят от разных параметров.

Состоятельность и эффективность в общем случае нескольких параметров

18.26 В общем случае, когда не обязательно имеется k достаточных статистик для k параметров, совместные МП-оценки обладают асимптотическими оптимальными свойствами, аналогичными тем, которые имели место в случае одного параметра.

Заметим, прежде всего, что доказательство состоятельности, данное в **18.10**, полностью переносится на случай многих параметров, если интерпретировать θ как вектор параметров $\theta_1, \ldots, \theta_k$ и $\hat{\theta}$, θ^* — как векторы оценок для θ. Таким образом, при очень общих условиях система совместных МП-оценок сходится по вероятности к истинной системе значений параметров θ_0.

Кроме того, непосредственно обобщая метод пункта **18.16**, можно показать (см., например, Вальд (1943а)), что совместное распределение МП-оценок сходится, в условиях регулярности, к многомерному нормальному распределению с матрицей рассеяния, обратной к матрице

$$(V_{rs}^{-1}) = -\mathbf{M}\left(\frac{\partial^2 \log L}{\partial\theta_r\,\partial\theta_s}\right) = \mathbf{M}\left(\frac{\partial \log L}{\partial\theta_r} \cdot \frac{\partial \log L}{\partial\theta_s}\right). \qquad (18.60)$$

Мы приведем лишь схему доказательства этого факта. Аналог разложения Тейлора (18.31), в котором левая часть приравнена нулю, принимает вид

$$\left(\frac{\partial \log L}{\partial \theta_r}\right)_{\theta_0} = \sum_{s=1}^{k} (\hat{\theta}_s - \theta_{s0}) \left(-\frac{\partial^2 \log L}{\partial \theta_r \partial \theta_s}\right)_{\theta^*}, \quad r = 1, 2, \ldots, k. \quad (18.61)$$

Поскольку θ^* сходится по вероятности к θ_0, а вторые производные в правой части (18.61) сходятся по вероятности к своим математическим ожиданиям, то мы можем рассматривать (18.61) как систему линейных уравнений относительно $(\hat{\theta}_r - \theta_{r0})$. Эту систему можно записать в виде

$$y = V^{-1}z, \quad (18.62)$$

где $y = \partial \log L/\partial\theta$, $z = \hat{\theta} - \theta_0$, а V^{-1} определена в (18.60).

Согласно многомерной центральной предельной теореме, если для каждого θ_r будет выполняться (18.29), то вектор y будет иметь предельное многомерное нормальное распределение. Следовательно, вектор z также будет в пределе распределен нормально. Матрица рассеяния V^{-1} вектора y дается, по определению, выражением (18.60), так что под знаком экспоненты многомерного распределения будет стоять квадратичная форма (см. **15.3**)

$$-\frac{1}{2} y'Vy. \quad (18.63)$$

Преобразование (18.62) дает для z квадратичную форму

$$-\frac{1}{2} z'V^{-1}z,$$

откуда следует, что матрица рассеяния z равна $(V^{-1})^{-1} = V$, как и утверждалось в (18.60).

18.27 Если имеется k совместно достаточных статистик для k параметров, то можно подставить (18.55) в (18.60), что дает следующее значение для матрицы, обратной к матрице рассеяния, в случае больших выборок:

$$(V_{rs}^{-1}) = -\left(\frac{\partial^2 \log L}{\partial \theta_r \partial \theta_s}\right)_{\hat{\theta}=\theta} = \left(\frac{\partial \log L}{\partial \theta_r} \cdot \frac{\partial \log L}{\partial \theta_s}\right)_{\hat{\theta}=\theta}. \quad (18.64)$$

Соотношение (18.64) так же, как и результат пункта **18.18**, обобщением которого оно является, устраняет необходимость нахождения математических ожиданий.

Если системы из k достаточных статистик не существует, то элементы матрицы рассеяния могут быть оценены по выборке стандартными методами.

18.28 Тот факт, что МП-оценки имеют матрицу рассеяния, определяемую соотношением (18.60), позволяет установить дальнейшие оптимальные свойства совместных МП-оценок.

Рассмотрим произвольную систему $t_1, \ldots, t_k$ состоятельных, функционально не связанных оценок параметров $\theta_1, \ldots, \theta_k$, имеющих матрицу рассеяния $\boldsymbol{D}$. Аналогично (17.21), при условиях регулярности, если каждая из состоятельных оценок является асимптотически несмещенной, получаем асимптотическое равенство

$$\int \ldots \int t_i L(x \mid \theta)\, dx_1 \ldots dx_n = \theta_i,$$

которое после дифференцирования приводит к соотношению

$$\int \ldots \int t_i \frac{\partial \log L}{\partial \theta_j} L\, dx_1 \ldots, dx_n = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

откуда, учитывая (17.18), получаем

$$\operatorname{cov}\left(t_i, \frac{\partial \log L}{\partial \theta_j}\right) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \tag{18.65}$$

Рассмотрим теперь матрицу рассеяния $2k$ вариант $t_1, \ldots, t_k, \frac{\partial \log L}{\partial \theta_1}, \ldots, \frac{\partial \log L}{\partial \theta_k}$, которая вследствие (18.65) и результатов **18.26** равна

$$\boldsymbol{C} = \begin{pmatrix} \boldsymbol{D} & \boldsymbol{I}_k \\ \boldsymbol{I}_k & \boldsymbol{V}^{-1} \end{pmatrix}, \tag{18.66}$$

где $\boldsymbol{I}_k$ — единичная матрица порядка k. Будучи матрицей рассеяния, $\boldsymbol{C}$ неотрицательно определена. Следовательно, ее определитель неотрицателен (см. (18.59)). Поскольку определитель матрицы

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{I}_k & -\boldsymbol{V} \\ \boldsymbol{0} & \boldsymbol{V} \end{pmatrix}$$

также неотрицателен, то

$$|\boldsymbol{MC}| = |\boldsymbol{M}|\,|\boldsymbol{C}| \geqslant 0.$$

Но так как

$$\boldsymbol{MC} = \begin{pmatrix} \boldsymbol{D} - \boldsymbol{V} & \boldsymbol{0} \\ \boldsymbol{V} & \boldsymbol{I}_k \end{pmatrix},$$

то $|\boldsymbol{D} - \boldsymbol{V}| \geqslant 0$, откуда (см. **19.8** ниже)

$$|\boldsymbol{D}| \geqslant |\boldsymbol{V}|. \tag{18.67}$$

Таким образом, определитель матрицы рассеяния любой системы оценок (называемый *обобщенной дисперсией* этой системы

оценок) асимптотически не меньше, чем $|\boldsymbol{V}|$. Но в **18.26** было показано, что для МП-оценок асимптотически $|\boldsymbol{D}| = |\boldsymbol{V}|$. Следовательно, МП-оценки являются асимптотически оценками с минимальной обобщенной дисперсией. Этот результат впервые был получен Гири (1942a).

Пример 18.13

Рассмотрим снова оценки $\bar{x}$ и s^2 примера 18.11. Имеем

$$\frac{\partial^2 \log L}{\partial \mu^2} = -\frac{n}{\sigma^2}, \quad \frac{\partial^2 \log L}{\partial (\sigma^2)^2} = \frac{n}{2\sigma^4} - \frac{\sum (x-\mu)^2}{\sigma^6},$$

$$\frac{\partial^2 \log L}{\partial \mu \, \partial (\sigma^2)} = -\frac{n(\bar{x}-\mu)}{\sigma^4}.$$

Вспоминая, что МП-оценки $\bar{x}$ и s^2 достаточны для μ и σ^2, и используя (18.64), мы получим матрицу, обратную к их асимптотической матрице рассеяния, если подставим во вторые производные значения $\bar{x} = \mu$ и $\sum (x-\mu)^2 = n\sigma^2$. Находим

$$\boldsymbol{V}^{-1} = \begin{pmatrix} n/\sigma^2 & 0 \\ 0 & n/2\sigma^4 \end{pmatrix},$$

откуда

$$\boldsymbol{V} = \begin{pmatrix} \sigma^2/n & 0 \\ 0 & 2\sigma^4/n \end{pmatrix}.$$

Мы видим, что $\bar{x}$ и s^2 асимптотически независимы и имеют нормальные распределения с указанными дисперсиями. При этом, как уже известно, свойство независимости, а также нормальность распределения $\bar{x}$ и выражение для ее дисперсии справедливы при любом n (примеры 11.3, 11.12). Что же касается s^2, то свойство нормальности и выражение для дисперсии этой оценки являются чисто предельными, поскольку (пример 11.7) ns^2/σ^2 распределено, как χ^2 с $n-1$ степенями свободы, и, следовательно, точная дисперсия s^2 (см. (16.5)) равна

$$\left(\frac{\sigma^2}{n}\right)^2 \cdot 2(n-1) = \frac{2\sigma^4(n-1)}{n^2}.$$

18.29 Может случиться, что некоторые из k параметров, от которых зависит распределение, известны и надо оценить остальные. При условиях регулярности МП-оценки нужных параметров получаются как решения соответствующей подсистемы системы уравнений правдоподобия (18.47). Очевидно, не следует ожидать, что на МП-оценку отдельного параметра не повлияет знание других параметров распределения. Вид МП-оценки некоторого параметра зависит от того, какие параметры оцениваются совместно с ним. Это хорошо видно на следующем примере.

Пример 18.14

В случае двумерного нормального распределения

$$dF(x, y) = \frac{dx\,dy}{2\pi\sigma_1\sigma_2(1-\rho^2)^{1/2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right\}\right],$$

$$-\infty \leqslant x,\ y \leqslant \infty;\quad \sigma_1,\ \sigma_2 > 0;\quad |\rho| < 1,$$

логарифм ФП равен

$$\log L(x, y \mid \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) =$$

$$= -n\log(2\pi) - \frac{1}{2}n\{\log\sigma_1^2 + \log\sigma_2^2 + \log(1-\rho^2)\} -$$

$$-\frac{1}{2(1-\rho^2)}\sum\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right\}.$$

Следовательно, пять уравнений правдоподобия имеют вид

$$\left.\begin{aligned}\frac{\partial \log L}{\partial\mu_1} &= \frac{n}{\sigma_1(1-\rho^2)}\left\{\frac{(\bar{x}-\mu_1)}{\sigma_1} - \rho\frac{(\bar{y}-\mu_2)}{\sigma_2}\right\} = 0,\\ \frac{\partial \log L}{\partial\mu_2} &= \frac{n}{\sigma_2(1-\rho^2)}\left\{\frac{(\bar{y}-\mu_2)}{\sigma_2} - \rho\frac{(\bar{x}-\mu_1)}{\sigma_1}\right\} = 0,\end{aligned}\right\} \quad (18.68)$$

$$\left.\begin{aligned}\frac{\partial \log L}{\partial(\sigma_1^2)} &= -\frac{1}{2\sigma_1^2(1-\rho^2)}\left\{n(1-\rho^2) - \frac{\sum(x-\mu_1)^2}{\sigma_1^2} + \rho\frac{\sum(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right\} = 0,\\ \frac{\partial \log L}{\partial(\sigma_2^2)} &= -\frac{1}{2\sigma_2^2(1-\rho^2)}\left\{n(1-\rho^2) - \frac{\sum(y-\mu_2)^2}{\sigma_2^2} + \rho\frac{\sum(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right\} = 0,\end{aligned}\right\} \quad (18.69)$$

$$\frac{\partial \log L}{\partial\rho} = \frac{1}{1-\rho^2}\left\{n\rho - \frac{1}{(1-\rho^2)}\left[\rho\left(\frac{\sum(x-\mu_1)^2}{\sigma_1^2} + \frac{\sum(y-\mu_2)^2}{\sigma_2^2}\right) - (1+\rho^2)\frac{\sum(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right]\right\} = 0. \quad (18.70)$$

(а) Предположим сначала, что надо оценить только ρ, а остальные четыре параметра известны. В этом случае надо решить одно уравнение (18.70). В нормированном виде мы уже имели дело с этой задачей в примере 18.3, где (18.70) приводит к кубическому уравнению для МП-оценки $\hat{\rho}$.

(б) Пусть теперь оцениваются σ_1^2, σ_2^2 и ρ, а μ_1 и μ_2 известны. Мы должны решить три уравнения правдоподобия (18.69) и (18.70). Вынося за фигурные скобки ненулевые сомножители, получаем после небольших преобразований

$$\left.\begin{aligned} n(1-\rho^2) &= \frac{\sum (x-\mu_1)^2}{\sigma_1^2} - \rho \frac{\sum (x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}, \\ n(1-\rho^2) &= \frac{\sum (y-\mu_2)^2}{\sigma_2^2} - \rho \frac{\sum (x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} \end{aligned}\right\} \tag{18.71}$$

и

$$n(1-\rho^2) = \frac{\sum (x-\mu_1)^2}{\sigma_1^2} + \frac{\sum (y-\mu_2)^2}{\sigma_2^2} - \frac{1+\rho^2}{\rho} \cdot \frac{\sum (x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}. \tag{18.72}$$

Если мы сложим уравнения (18.71) и вычтем из суммы уравнение (18.72), то получим

$$n(1-\rho^2) = \frac{1-\rho^2}{\rho} \frac{\sum (x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2},$$

или

$$\rho = \frac{\frac{1}{n}\sum (x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} \tag{18.73}$$

Подстановка (18.73) в (18.71) дает

$$\hat{\sigma}_1^2 = \frac{1}{n}\sum (x-\mu_1)^2, \quad \hat{\sigma}_2^2 = \frac{1}{n}\sum (y-\mu_2)^2. \tag{18.74}$$

Следовательно, согласно (18.73)

$$\hat{\rho} = \frac{\frac{1}{n}\sum (x-\mu_1)(y-\mu_2)}{\hat{\sigma}_1\hat{\sigma}_2}. \tag{18.75}$$

Таким образом, в этом случае МП-оценка $\hat{\rho}$ есть просто выборочный коэффициент корреляции, вычисленный относительно известных генеральных средних.

(в) Предположим, наконец, что требуется оценить все пять параметров распределения. Тогда надо решать совместно пять уравнений (18.68), (18.69) и (18.70). (18.68) сводится к системе

$$\left.\begin{aligned} \frac{(\bar{x}-\mu_1)}{\sigma_1} &= \rho \frac{(\bar{y}-\mu_2)}{\sigma_2}, \\ \frac{(\bar{y}-\mu_2)}{\sigma_2} &= \rho \frac{(\bar{x}-\mu_1)}{\sigma_1}, \end{aligned}\right\} \tag{18.76}$$

которая имеет единственное решение

$$\bar{x} - \mu_1 = \bar{y} - \mu_2 = 0. \tag{18.77}$$

Вместе с (18.74) и (18.75), представляющими собой решения уравнений (18.69) и (18.70), (18.77) дает систему из пяти МП-оценок:

$$\left.\begin{aligned} \hat{\mu}_1 &= \bar{x}, \quad \hat{\sigma}_1^2 = \frac{1}{n}\sum (x-\bar{x})^2, \\ \hat{\mu}_2 &= \bar{y}, \quad \hat{\sigma}_2^2 = \frac{1}{n}\sum (y-\bar{y})^2, \\ \hat{\rho} &= \frac{\frac{1}{n}\sum (x-\bar{x})(y-\bar{y})}{\hat{\sigma}_1\hat{\sigma}_2}. \end{aligned}\right\} \tag{18.78}$$

Таким образом, МП-оценками всех пяти параметров служат соответствующие выборочные средние.

18.30 Поскольку МП-оценки некоторого параметра распределения будут различными функциями наблюдений в зависимости от того, какие из других параметров этого распределения известны, то асимптотические дисперсии оценок будут также разными. Чтобы облегчить вычисление матриц рассеяния МП-оценок в случае, когда распределение допускает систему из k достаточных статистик для k параметров, можно воспользоваться представлением (18.64) для матрицы, обратной к асимптотической матрице рассеяния МП-оценок.

Пример 18.15

Вычислим асимптотические матрицы рассеяния МП-оценок для каждого из трех случаев, рассмотренных в примере 18.14.

(а) Когда мы оцениваем только ρ, $\hat{\rho}$ не будет достаточной статистикой. Однако мы уже вычисляли асимптотическую дисперсию этой оценки в примере 18.6 и нашли, что

$$\mathrm{D}\hat{\rho} = \frac{(1-\rho^2)^2}{n(1+\rho^2)}. \tag{18.79}$$

Тот факт, что мы рассматривали нормированное генеральное распределение, не имеет значения, поскольку коэффициент корреляции ρ инвариантен по отношению к изменениям начала отсчета и масштаба.

(б) В случае оценивания трех параметров σ_1^2, σ_2^2 и ρ три МП-оценки, даваемые выражениями (18.74) и (18.75), совместно достаточны. Следовательно, мы можем использовать

(18.64). Записывая параметры в указанном выше порядке, находим, что матрица 3×3, обратная к матрице рассеяния соответствующих оценок, имеет вид

$$\boldsymbol{V}_3^{-1} = (V_{rs}^{-1}) = -\left(\frac{\partial^2 \log L}{\partial\theta_r\,\partial\theta_s}\right)_{\hat{\theta}=\theta} =$$

$$= \frac{n}{1-\rho^2}\begin{pmatrix} \dfrac{2-\rho^2}{4\sigma_1^4} & \dfrac{-\rho^2}{4\sigma_1^2\sigma_2^2} & \dfrac{-\rho}{2\sigma_1^2} \\ \dfrac{-\rho^2}{4\sigma_1^2\sigma_2^2} & \dfrac{2-\rho^2}{4\sigma_2^4} & \dfrac{-\rho}{2\sigma_2^2} \\ \dfrac{-\rho}{2\sigma_1^2} & \dfrac{-\rho}{2\sigma_2^2} & \dfrac{1+\rho^2}{1-\rho^2} \end{pmatrix}. \qquad (18.80)$$

Обращение (18.80) дает асимптотическую матрицу рассеяния

$$\boldsymbol{V}_3 = \frac{1}{n}\begin{pmatrix} 2\sigma_1^4 & 2\rho^2\sigma_1^2\sigma_2^2 & \rho(1-\rho^2)\sigma_1^2 \\ 2\rho^2\sigma_1^2\sigma_2^2 & 2\sigma_2^4 & \rho(1-\rho^2)\sigma_2^2 \\ \rho(1-\rho^2)\sigma_1^2 & \rho(1-\rho^2)\sigma_2^2 & (1-\rho^2)^2 \end{pmatrix}. \qquad (18.81)$$

(в) При оценивании всех пяти параметров достаточную систему образуют МП-оценки (18.78). Кроме того, (3×3)-матрица $\boldsymbol{V}_3^{-1}$ из (18.80) является частью (5×5)-матрицы $\boldsymbol{V}_5^{-1}$, обратной к матрице рассеяния, которую мы ищем. Если записать параметры в порядке μ_1, μ_2, σ_1^2, σ_2^2, ρ, то (18.80) будет нижним главным минором 3×3 для $\boldsymbol{V}_5^{-1}$. Для элементов, включающих производные по μ_1 и μ_2, согласно (18.68) получаются следующие выражения:

$$\left.\begin{aligned} \frac{\partial^2 \log L}{\partial\mu_1^2} &= \frac{-n}{\sigma_1^2(1-\rho^2)}, \qquad \frac{\partial^2 \log L}{\partial\mu_1\,\partial\mu_2} = \frac{n\rho}{\sigma_1\sigma_2(1-\rho^2)}, \\ \frac{\partial^2 \log L}{\partial\mu_2^2} &= \frac{-n}{\sigma_2^2(1-\rho^2)}. \end{aligned}\right\} \qquad (18.82)$$

При $\tilde{x} = \mu_1$ и $\tilde{y} = \mu_2$ имеем

$$\frac{\partial^2 \log L}{\partial\mu_i\,\partial\sigma_1^2} = \frac{\partial^2 \log L}{\partial\mu_i\,\partial\sigma_2^2} = \frac{\partial^2 \log L}{\partial\mu_i\,\partial\rho} = 0, \qquad i = 1,\ 2. \qquad (18.83)$$

Обозначая теперь $\boldsymbol{V}_2^{-1}$ матрицу, обратную к матрице рассеяния МП-оценок для μ_1 и μ_2, равную

$$\boldsymbol{V}_2^{-1} = \frac{n}{1-\rho^2}\begin{pmatrix} \dfrac{1}{\sigma_1^2} & \dfrac{-\rho}{\sigma_1\sigma_2} \\ \dfrac{-\rho}{\sigma_1\sigma_2} & \dfrac{1}{\sigma_2^2} \end{pmatrix}, \qquad (18.84)$$

получаем

$$\boldsymbol{V}_5^{-1} = \begin{pmatrix} \boldsymbol{V}_2^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{V}_3^{-1} \end{pmatrix}. \tag{18.85}$$

Чтобы получить ненулевые элементы матрицы $\boldsymbol{V}_5$, можно обращать $\boldsymbol{V}_2^{-1}$ и $\boldsymbol{V}_3^{-1}$ отдельно. Результатом обращения $\boldsymbol{V}_3^{-1}$ является (18.81). Обращение $\boldsymbol{V}_2^{-1}$ приводит к матрице

$$\boldsymbol{V}_2 = \frac{1}{n} \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}. \tag{18.86}$$

Таким образом,

$$\boldsymbol{V}_5 = \begin{pmatrix} \boldsymbol{V}_2 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{V}_3 \end{pmatrix}, \tag{18.87}$$

где $\boldsymbol{V}_2$ и $\boldsymbol{V}_3$ определены равенствами (18.86) и (18.81).

Из полученного результата следует, что (ранее это было показано (см. **16.25**) для произвольных объемов выборок) выборочные средние в двумерной нормальной выборке распределены независимо от выборочных дисперсий и ковариаций и что коэффициент корреляции между выборочными средними равен ρ, а между выборочными дисперсиями равен ρ^2 (пример 13.1).

Кейл (1962) рассматривает различные итерационные методы решения уравнений правдоподобия для нескольких параметров. $\boldsymbol{\theta}_0$ в (18.62) заменяется пробным вектором $\boldsymbol{t}$ (см. **18.21**), так что $\hat{\boldsymbol{\theta}} = \boldsymbol{t} + \boldsymbol{V}\boldsymbol{y}$ может быть итерировано столько раз, сколько необходимо.

Неодинаковые исходные распределения

18.31 Мы уже довольно подробно изучили метод максимума правдоподобия, но касались пока лишь задач оценивания для наблюдений, полученных из одного и того же распределения. Теперь мы кратко рассмотрим поведение МП-оценок в случае, когда это условие не выполняется. Для этого мы должны заменить (18.1) более общей ФП

$$L(x \mid \theta_1, \ldots, \theta_k) = f_1(x_1 \mid \theta) f_2(x_2 \mid \theta) \ldots f_n(x_n \mid \theta), \tag{18.88}$$

где различные сомножители f_i могут зависеть от различных функций от системы параметров $\theta_1, \ldots, \theta_k$.

В этой ситуации МП-оценки могут даже не быть состоятельными, а для одного класса случаев, когда число параметров возрастает с ростом числа наблюдений (k зависит от n), метод максимума правдоподобия может стать неэффективным. Следующие два примера служат этому иллюстрацией.

Пример 18.16

Пусть x_i — нормальная варианта со средним θ_i и дисперсией $\sigma^2 > 0$ $(i = 1, 2, \ldots, n)$. Функция правдоподобия (18.88) имеет вид

$$\log L = -\frac{1}{2} n \log (2\pi) - \frac{1}{2} n \log (\sigma^2) - \frac{1}{2\sigma^2} \sum_i (x_i - \theta_i)^2,$$

а уравнение

$$\frac{\partial \log L}{\partial (\sigma^2)} = -\frac{n}{2\sigma^2} + \frac{\sum_i (x_i - \theta_i)^2}{2\sigma^4} = 0$$

дает МП-оценку

$$\hat{\sigma}^2 = \frac{1}{n} \sum_i (x_i - \theta_i)^2. \tag{18.89}$$

Поскольку, однако, мы имеем только по одному наблюдению над каждым из различных нормальных распределений, то получаем также

$$\hat{\theta}_i = x_i. \tag{18.90}$$

Таким образом, при совместном оценивании σ^2 и θ_i $(i = 1, 2, \ldots, n)$ оценки (18.89) и (18.90) приводят к абсурдному результату:

$$\hat{\sigma}^2 = 0.$$

Мы не можем надеяться получить эффективной оценки для σ^2, имея по одному наблюдению из каждого распределения: (18.90) — совершенно бесполезная оценка.

Ситуация ненамного улучшится, если мы будем иметь по два наблюдения из каждого нормального распределения. В этом случае получим

$$\left.\begin{aligned} \hat{\theta}_i &= \frac{1}{2}(x_{i1} + x_{i2}) = \bar{x}_i, \\ \hat{\sigma}^2 &= \frac{1}{2n} \sum_{i=1}^{n} \sum_{j=1}^{2} (x_{ij} - \bar{x}_i)^2. \end{aligned}\right\} \tag{18.91}$$

Но так как

$$\mathsf{M}\left\{\frac{1}{2} \sum_{j=1}^{2} (x_{ij} - \bar{x}_i)^2\right\} = \frac{1}{2}\sigma^2$$

(ср. пример 17.3), то из (18.91) следует, что для всех n

$$\mathsf{M}(\hat{\sigma}^2) = \frac{1}{2}\sigma^2,$$

т. е. $\hat{\sigma}^2$ не состоятельна (см. Нейман и Скотт (1948)). Дело в том, что смещение МП-оценок при малых выборках (18.14) сохраняется в этом примере при возрастании n в силу того, что число распределений также растет с ростом n.

Другие примеры подобного типа рассматривались в литературе в связи с приложениями, в которых они возникали. Нам они также встретятся в более поздних главах, в частности в главе 29. Сейчас же необходимо лишь подчеркнуть, что в нестандартных ситуациях требуется тщательное изучение свойств МП-оценок. Не следует предполагать, что асимптотические оптимальные свойства сохраняются во всех случаях. Например, МП-оценка может даже не существовать (см. упражнение 18.34) или может иметься несколько МП-оценок, некоторые из которых будут несостоятельными (см. упражнение 18.35).

Использование функции правдоподобия

18.32 Как это рекомендует Фишер (1956), всегда можно изучить поведение ФП во всей области изменения θ и построить график ФП. Хотя это и дает общую информацию, по-видимому, таким образом нельзя получить выводов, представляющих непосредственную ценность для задачи оценивания. ФП содержит всю информацию, имеющуюся в выборке, строго говоря, в том смысле, что, как было замечено в **17.38**, сами наблюдения являются системой совместно достаточных статистик для любой задачи. Такой подход обладает тем достоинством, что обращает внимание на необходимость выбора функциональной формы распределения (или распределений), порождающего наблюдения, до того, как ФП вообще может быть использована для получения МП-оценок или для других целей. Другими словами, некоторая информация (в общем смысле) должна быть дана статистиком. Если же он не может или не хочет дать эту информацию, то следует обратиться к непараметрическим гипотезам, которые приводят к совершенно другим методам. Эти методы будут рассмотрены позже.

В **23.37** мы рассмотрим некоторые недавние результаты, связывающие использование ФП с условными статистическими выводами.

Оценка параметров сдвига и масштаба

18.33 Следуя Фишеру (1921а), можно использовать метод максимума правдоподобия для нахождения эффективных оценок параметров сдвига и масштаба для распределения любой заданной формы.

Рассмотрим функцию плотности вида

$$dF(x) = f\left(\frac{x-\alpha}{\beta}\right) d\left(\frac{x-\alpha}{\beta}\right), \quad \beta > 0. \tag{18.92}$$

Параметр α называется параметром расположения (или параметром сдвига), а β — параметром масштаба. Перепишем (18.92) в виде

$$dF = \exp\{g(y)\}\, dy = \exp\{g(y)\}\, dx/\beta, \tag{18.93}$$

где

$$y = (x-\alpha)/\beta, \quad g(y) = \log f(y).$$

ФП для выборки объема n имеет вид

$$\log L(x \mid \alpha, \beta) = \sum_{i=1}^{n} g(y_i) - n \log \beta. \tag{18.94}$$

Отсюда получаем уравнения правдоподобия

$$\left.\begin{aligned} \frac{\partial \log L}{\partial \alpha} &= -\frac{1}{\beta}\sum_i g'(y_i) = 0, \\ \frac{\partial \log L}{\partial \beta} &= -\frac{1}{\beta}\left\{\sum_i y_i g'(y_i) + n\right\} = 0, \end{aligned}\right\} \tag{18.95}$$

где $g'(y) = \partial g(y)/\partial y$. В условиях регулярности решение (18.95) дает МП-оценки $\hat{\alpha}$ и $\hat{\beta}$. Предположим теперь, что для всех допустимых значений α и β выполняются соотношения

$$\left.\begin{aligned} \mathbf{M}\left(\frac{\partial \log L}{\partial \alpha}\right) &= -\frac{n}{\beta}\mathbf{M}\{g'(y)\} = 0, \\ \mathbf{M}\left(\frac{\partial \log L}{\partial \beta}\right) &= -\frac{n}{\beta}[\mathbf{M}\{y g'(y)\} + 1] = 0. \end{aligned}\right\} \tag{18.96}$$

Как и в случае (17.18), соотношения (18.96) выполняются, если можно в левой части дифференцировать под знаком интеграла. Можно переписать (18.96) следующим образом:

$$\mathbf{M}\{g'(y)\} = 0, \tag{18.97}$$

$$\mathbf{M}\{y g'(y)\} = -1. \tag{18.98}$$

Вычислим теперь элементы матрицы (18.60), обратной к матрице рассеяния, считая справедливым соотношение

$$\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta_r\, \partial \theta_s}\right) = -\mathbf{M}\left(\frac{\partial \log L}{\partial \theta_r} \cdot \frac{\partial \log L}{\partial \theta_s}\right).$$

Из (18.95) следует (аргумент у $g(y)$ опущен)

$$\left.\begin{aligned} \mathbf{M}\left(\frac{\partial^2 \log L}{\partial \alpha^2}\right) &= \frac{n}{\beta^2}\,\mathbf{M}\,(g''), \\ \mathbf{M}\left(\frac{\partial^2 \log L}{\partial \beta^2}\right) &= \frac{n}{\beta^2}\,\mathbf{M}\,(g''y^2 + 2g'y + 1). \end{aligned}\right\} \tag{18.99}$$

Отсюда, учитывая (18.98), получаем

$$\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \beta^2}\right) = \frac{n}{\beta^2}\,\mathbf{M}\,(g''y^2 - 1). \tag{18.100}$$

Кроме того,

$$\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \alpha\, \partial \beta}\right) = \frac{n}{\beta^2}\,\mathbf{M}\,(g' + g''y),$$

или, вследствие (18.97),

$$\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \alpha\, \partial \beta}\right) = \frac{n}{\beta^2}\,\mathbf{M}\,(g''y). \tag{18.101}$$

Равенства (18.99)—(18.101) приводят к следующему выражению для матрицы (18.60):

$$\boldsymbol{V}^{-1} = -\frac{n}{\beta^2}\,\mathbf{M}\begin{pmatrix} g'' & g''y \\ g''y & g''y^2 - 1 \end{pmatrix}. \tag{18.102}$$

Дисперсии и ковариации получаются обращением этой матрицы. Конечно, если α и β оценивается отдельно, то дисперсия МП-оценки будет равна обратной величине соответствующего элемента главной диагонали матрицы $\boldsymbol{V}^{-1}$.

18.34 Если $g(y)$ является четной функцией своего аргумента, т. е. распределение симметрично относительно α, то (18.102) упрощается. Действительно, в этом случае имеем

$$g(y) = g(-y), \quad g'(y) = -g'(-y), \quad g''(y) = g''(-y), \tag{18.103}$$

откуда следует, что внедиагональные элементы матрицы (18.102) равны нулю:

$$\mathbf{M}\{g''(y)\,y\} = 0. \tag{18.104}$$

Таким образом, для симметричных распределений матрица (18.102) будет диагональной. Следовательно, МП-оценки параметров сдвига и масштаба любого симметричного распределения, удовлетворяющего нашим условиям регулярности, будут асимптотически некоррелированы и (поскольку они согласно **18.26** имеют асимптотически двумерное нормальное распределение) асимптотически независимы. Это относится, в частности, к нормальному распределению, для которого этот результат был получен в примере 18.13.

18.35 В случае несимметричных распределений внедиагональные элементы в (18.102) можно обратить в нуль простым изменением начала отсчета. Положим

$$z = y - \frac{\mathbf{M}(g''y)}{\mathbf{M}(g'')}. \tag{18.105}$$

Тогда

$$\mathbf{M}(g''y) = \mathbf{M}\left\{ g'' \left[z + \frac{\mathbf{M}(g''y)}{\mathbf{M}(g'')} \right] \right\} = \mathbf{M}(g''z) + \mathbf{M}(g''y),$$

откуда

$$\mathbf{M}(g''z) = 0. \tag{18.106}$$

Таким образом, если начало отсчета определяется соотношением (18.105), то (18.102) становится диагональной матрицей и дисперсии оценок определяются просто как обратные величины диагональных элементов. Начало отсчета, при котором оценки некоррелированы, называется *центром расположения* распределения. Причиной выбора центра расположения в качестве начала отсчета служит то, что в случае применения итерационной процедуры оценки могут вычисляться отдельно.

Пример 18.17

Распределение

$$dF(x) = \frac{1}{\Gamma(p)} \left(\frac{x-\alpha}{\beta} \right)^{p-1} \exp\left\{ -\frac{(x-\alpha)}{\beta} \right\} d\left(\frac{x-\alpha}{\beta} \right), \quad \alpha \leqslant x \leqslant \infty,$$
$$\beta > 0, \quad p > 2,$$

имеет область определения, зависящую от α. Однако это распределение и его первая производная по α равны нулю при $x = \alpha$ для $p > 2$ (см. упражнение 17.23), поэтому наши условия регулярности выполнены. В данном случае имеем

$$g(y) = -\log \Gamma(p) + (p-1)\log y - y,$$
$$\mathbf{M}(g'') = \mathbf{M}\left\{ -\frac{p-1}{y^2} \right\} = -\frac{1}{p-2},$$
$$\mathbf{M}(g''y) = \mathbf{M}\left\{ -\frac{p-1}{y} \right\} = -1,$$
$$\mathbf{M}(g''y^2) = \mathbf{M}\{-(p-1)\} = -(p-1).$$

Следовательно, центр расположения, определяемый соотношением (18.105), равен

$$\frac{\mathbf{M}(g''y)}{\mathbf{M}(g'')} = p - 2.$$

Матрица (18.102) имеет вид

$$\boldsymbol{V}^{-1} = \frac{n}{\beta^2}\begin{pmatrix} \frac{1}{p-2} & 1 \\ 1 & p \end{pmatrix}, \tag{18.107}$$

откуда легко получить обратную ей матрицу рассеяния оценок $\hat{\alpha}$ и $\hat{\beta}$:

$$\boldsymbol{V} = \frac{(p-2)\beta^2}{2n}\begin{pmatrix} p & -1 \\ -1 & \frac{1}{p-2} \end{pmatrix}. \tag{18.108}$$

Если начало отсчета совпадает с центром расположения, то для некоррелированных оценок $\hat{\alpha}_u$ и $\hat{\beta}_u$ получим

$$\mathrm{D}\hat{\alpha}_u = (p-2)\beta^2/n, \qquad \mathrm{D}\hat{\beta}_u = \beta^2/(2n). \tag{18.109}$$

Сравнивая (18.109) с (18.108) и (18.107), мы видим, что $\mathrm{D}\hat{\beta}$ не изменяется при изменении начала отсчета, а $\mathrm{D}\hat{\alpha}_u$ совпадает с $\mathrm{D}\hat{\alpha}$, вычисленной для случая, когда оценивается один параметр α.

Эффективность метода моментов

18.36 В главе 6 были рассмотрены распределения Пирсона. Однако в основном там изучались свойства самих генеральных совокупностей и не ставилось никаких вопросов относительно надежности оценок. Предположим теперь, что имеется выборка из генеральной совокупности. Обеспечивает ли подгонка моментов наиболее эффективные оценки неизвестных параметров? Сейчас мы увидим, что в общем случае это не так.

Пусть распределение зависит от четырех параметров. Для того чтобы МП-оценки этих параметров выражались в виде линейных функций от моментов (как при подгонке кривых Пирсона), мы должны иметь

$$\frac{\partial \log L}{\partial \theta_r} = a_0 + a_1 \sum x + a_2 \sum x^2 + a_3 \sum x^3 + a_4 \sum x^4, \tag{18.110}$$

$$r = 1, \ldots, 4,$$

откуда

$$f(x \mid \theta_1, \ldots, \theta_4) = \exp(b_0 + b_1 x + b_2 x^2 + b_3 x^3 + b_4 x^4), \tag{18.111}$$

где коэффициенты b зависят от параметров θ. Это наиболее общее представление распределений, для которых метод моментов дает МП-оценки. Коэффициенты b, конечно, подчинены условию, что интеграл от функции плотности сходится и равен единице.

Без потери общности можно положить $b_1 = 0$. Если, кроме того, b_3 и b_4 равны нулю, то распределение нормально и метод

моментов эффективен. В других случаях (18.111) приводит к распределению Пирсона только приближенно. Пусть, например,

$$\frac{\partial \log f}{\partial x} = 2b_2x + 3b_3x^2 + 4b_4x^3.$$

При малых b_3 и b_4 отсюда приближенно получается равенство

$$\frac{\partial \log f}{\partial x} = \frac{2b_2x}{1 - \frac{3b_3}{2b_2}x - \frac{2b_4}{b_2}x^2}, \tag{18.112}$$

являющееся одной из форм уравнения, определяющего распределения Пирсона (см. (6.1)). На высокую эффективность метода моментов можно надеяться лишь при малых по сравнению с b_2 коэффициентах b_3 и b_4.

18.37 Детальное исследование эффективности моментов при оценке параметров распределения Пирсона было проведено Фишером (1921a). Мы приводим здесь лишь один результат в качестве иллюстрирующего примера.

Пример 18.18

Рассмотрим гамма-распределение с тремя параметрами α, σ и p:

$$dF = \frac{1}{\sigma\Gamma(p)}\left(\frac{x-\alpha}{\sigma}\right)^{p-1}\exp\left\{-\left(\frac{x-\alpha}{\sigma}\right)\right\}dx, \quad \alpha \leqslant x \leqslant \infty;$$
$$\alpha > 0; \quad p > 2.$$

Функция правдоподобия записывается в виде

$$\log L = -np\log\sigma - n\log\Gamma(p) + (p-1)\sum\log(x-\alpha) - \sum(x-\alpha)/\sigma,$$

откуда получаем три уравнения правдоподобия:

$$\frac{\partial \log L}{\partial \alpha} = -(p-1)\sum\frac{1}{x-\alpha} + \frac{n}{\sigma} = 0,$$

$$\frac{\partial \log L}{\partial \sigma} = -\frac{np}{\sigma} + \frac{\sum(x-\alpha)}{\sigma^2} = 0,$$

$$\frac{\partial \log L}{\partial p} = -n\log\sigma - n\frac{d}{\alpha p}\log\Gamma(p) + \sum\log(x-\alpha) = 0.$$

Матрица (18.60), обратная к матрице рассеяния параметров α, σ и p, имеет вид

$$\boldsymbol{V}^{-1} = n\begin{Vmatrix} \frac{1}{\sigma^2(p-2)} & \frac{1}{\sigma^2} & \frac{1}{\sigma(p-1)} \\ \frac{1}{\sigma^2} & \frac{p}{\sigma^2} & \frac{1}{\sigma} \\ \frac{1}{\sigma(p-1)} & \frac{1}{\sigma} & \frac{d^2\log\Gamma(p)}{dp^2} \end{Vmatrix}.$$

Определитель этой матрицы равен

$$\frac{|V^{-1}|}{n^3} = \Delta = \frac{1}{(p-2)\sigma^4}\left\{2\frac{d^2\log\Gamma(p)}{dp^2} - \frac{2}{p-1} + \frac{1}{(p-1)^2}\right\}.$$

Отсюда находим дисперсии оценок:

$$\left.\begin{aligned}
\mathrm{D}\hat{\alpha} &= \frac{1}{n\Delta\sigma^2}\left\{p\frac{d^2\log\Gamma(p)}{dp^2} - 1\right\},\\
\mathrm{D}\hat{\sigma} &= \frac{1}{n\Delta\sigma^2}\left\{\frac{1}{p-2}\frac{d^2\log\Gamma(p)}{dp^2} - \frac{1}{(p-1)^2}\right\},\\
\mathrm{D}\hat{p} &= \frac{2}{n\Delta(p-2)\sigma^4} = \frac{2}{n}\Big/\left\{\frac{2d^2\log\Gamma(p)}{dp^2} - \frac{2}{p-1} + \frac{1}{(p-1)^2}\right\}.
\end{aligned}\right\} \quad (18.113)$$

Используя при больших p разложение Стирлинга, получаем

$$\frac{d^2}{dp^2}\log\Gamma(1+p) = \\ = \frac{d^2}{dp^2}\left\{\frac{1}{2}\log(2\pi) + \left(p+\frac{1}{2}\right)\log p - p + \frac{1}{12p} - \frac{1}{360p^3} + \dots\right\},$$

откуда

$$2\frac{d^2}{dp^2}\log\Gamma(1+p) - \frac{2}{p} + \frac{1}{p^2} = \frac{1}{3}\left\{\frac{1}{p^3} - \frac{1}{5p^5} + \frac{1}{7p^7} - \dots\right\}.$$

Следовательно, ввиду (18.113) имеем приближенно

$$\mathrm{D}\hat{p} \approx \frac{6}{n}\left\{(p-1)^3 + \frac{1}{5}(p-1)\right\}. \quad (18.114)$$

Если мы оцениваем параметры, приравнивая выборочные моменты и соответствующие моменты, выраженные через параметры, то получаем

$$\alpha + \sigma p = m_1, \quad \sigma^2 p = m_2, \quad 2\sigma^3 p = m_3,$$

так что при любых α и σ

$$b_1 = m_3^2/m_2^3 = 4/p, \quad (18.115)$$

где b_1 — выборочное значение коэффициента асимметрии β_1. Имеем (см. **10.15**)

$$\mathrm{D}b_1 = \frac{\beta_1}{n}\{4\beta_4 - 24\beta_2 + 36 + 9\beta_1\beta_2 - 12\beta_3 + 35\beta_1\}.$$

Для рассматриваемого распределения это выражение принимает вид

$$\mathrm{D}b_1 = \frac{\beta_1^2}{n}\cdot\frac{6(p+1)(p+5)}{p}. \quad (18.116)$$

Следовательно, учитывая (18.115), получаем для оценки $\tilde{p}$,

полученной по методу моментов,

$$\mathrm{D}\tilde{p} \approx \frac{p^4}{16}\,\mathrm{D}b_1 = \frac{6}{n}\,p(p+1)(p+5).$$

Для больших p эффективность этой оценки равна согласно (18.114)

$$\frac{\mathrm{D}\hat{p}}{\mathrm{D}\tilde{p}} = \frac{\left\{(p-1)^3 + \frac{1}{5}(p-1)\right\}}{p(p+1)(p+5)},$$

что, очевидно, меньше 1. Когда p превосходит 39,1 ($\beta_1 = 0{,}102$), эффективность превышает 80%. Для $p = 20$ ($\beta_1 = 0{,}20$) она равна 65%. Для $p = 5$ более точное вычисление, основанное на таблицах тригамма-функции $\frac{d^2 \log \Gamma(1+p)}{\partial p^2}$, показывает, что эффективность равна всего лишь 22%.

УПРАЖНЕНИЯ

18.1 В примере 18.7 показать, рассматривая случай $n = 1$, что МГД для выборок малого объема не достигается. Показать, что для рассматриваемого распределения эффективность выборочного среднего относительно МП-оценки равна 1/2.

18.2 Показать, что если МП-оценка $\hat{\theta}$ является корнем уравнения $\partial \log L/\partial\theta = 0$, то наиболее общий вид дифференцируемого по θ распределения, для которого $\hat{\theta} = \bar{x}$ (выборочное арифметическое среднее), есть

$$f(x \mid \theta) = \exp\{A(\theta) + A'(\theta)(x - \theta) + B(x)\}$$

и, следовательно, статистика $\bar{x}$ достаточна для θ, а соответствующая МГД равна $\{nA''(\theta)\}^{-1}$. Показать также, что если θ — параметр расположения, то f — нормальное распределение со средним θ (результат, принадлежащий Гауссу), тогда как если θ — параметр масштаба, то $f = \theta^{-1}\exp(-x/\theta)$.

(См. Кейнес, 1911, и Тейчер, 1961.)

18.3 Показать, что наиболее общее выражение для плотности непрерывного распределения, имеющего в качестве МП-оценки для параметра θ выборочное геометрическое среднее, имеет вид

$$f(x \mid \theta) = \left(\frac{x}{\theta}\right)^{\theta A'(\theta)} \exp\{A(\theta) + B(x)\}.$$

Показать далее, что гармоническому среднему соответствует следующее общее выражение для плотности распределения:

$$f(x \mid \theta) = \exp\left[\frac{1}{x}\{\theta A'(\theta) - A(\theta)\} - A'(\theta) + B(x)\right].$$

(Кейнес, 1911.)

18.4 Показать, что в каждом из случаев упражнения 18.3 МП-оценка достаточна для θ, но не является МГД-оценкой, в отличие от случая арифметического среднего, рассмотренного в примере 18.2. Найти в каждом случае функцию от θ, оцениваемую с дисперсией, равной соответствующей МГД, и вычислить эту МГД.

18.5 Показать, что для распределения

$$dF(x) \propto \exp\{-(x-\alpha)/\beta\}\,dx, \qquad 0 \leqslant \alpha \leqslant x \leqslant \beta,$$

$x_{(1)}$ и $x_{(n)}$ являются соответственно МП-оценками для α и β, но не образуют достаточную пару для α и β.

18.6 Показать, что в случае выборок объема n из распределения крайних значений (см. (14.66))

$$dF(x) = \alpha \exp\{-\alpha(x-\mu) - \exp[-\alpha(x-\mu)]\}\,dx, \qquad -\infty \leqslant x \leqslant \infty,$$

МП-оценки $\hat{\alpha}$ и $\hat{\mu}$ даются системой уравнений

$$\frac{1}{\hat{\alpha}} = \bar{x} - \frac{\sum x e^{-\hat{\alpha}x}}{\sum e^{-\hat{\alpha}x}},$$

$$e^{-\hat{\alpha}\hat{\mu}} = \frac{1}{n}\sum e^{-\hat{\alpha}x}$$

и что для больших выборок

$$\mathbf{D}\hat{\alpha} = \alpha^2/(\pi^2/6), \qquad \mathbf{D}\hat{\mu} = \frac{1}{\alpha^2}\left\{1 + \frac{(1-\gamma)^2}{\pi^2/6}\right\},$$

$$\operatorname{cov}(\hat{\alpha}, \hat{\mu}) = -(1-\gamma)/(\pi^2/6),$$

где γ — константа Эйлера, равная 0,5772...

(Б. Кимбелл, 1946.)

18.7 Если величина x распределена по нормальному закону

$$dF(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} dx, \qquad -\infty \leqslant x \leqslant \infty,$$

то логнормальное распределение величины $y = e^x$ имеет среднее $\theta_1 = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$ и дисперсию $\theta_2 = \exp(2\mu + \sigma^2)\{\exp(\sigma^2) - 1\}$ (см. (6.63)).

Показать, что МП-оценка для θ_1 равна

$$\hat{\theta}_1 = \exp\left(\bar{x} + \frac{1}{2}s^2\right),$$

где $\bar{x}$ и s^2 — выборочное среднее и выборочная дисперсия величины x, и что

$$\mathbf{M}(\hat{\theta}_1) = \mathbf{M}\{\exp(\bar{x})\}\,\mathbf{M}\left\{\exp\left(\frac{1}{2}s^2\right)\right\} =$$

$$= \theta_1 \exp\left\{-\frac{(n-1)\sigma^2}{2n}\right\}\left(1 - \frac{\sigma^2}{n}\right)^{-\frac{1}{2}(n-1)} > \theta_1,$$

т. е. оценка $\hat{\theta}_1$ смещена вверх. Показать также, что $\mathbf{M}(\hat{\theta}_1) \to \theta_1$ при $n \to \infty$ т. е. $\hat{\theta}_1$ асимптотически несмещена.

18.8 Пусть $f(t)$ определяется как сумма ряда

$$f(t) = 1 + t + \frac{n-1}{n+1}\frac{t^2}{2!} + \frac{(n-1)^2}{(n+1)(n+3)}\frac{t^3}{3!} + \ldots$$

Показать, что модифицированная МП-оценка для параметра θ_1 в упражнении 18.7

$$\bar{\theta}_1 = \exp(\bar{x})\, f\left(\frac{1}{2}s^2\right)$$

является несмещенной. Показать далее, что $\hat{\theta}_1 > \bar{\theta}_1$ для всех выборок, т. е. $\hat{\theta}_1$ равномерно смещена относительно $\bar{\theta}_1$.

18.9 Показать, что точная дисперсия оценки $\hat{\theta}_1$, рассматриваемой в упражнении 18.7, равна

$$\mathbf{D}\hat{\theta}_1 = \mathbf{M}\{\exp(2\bar{x})\}\,\mathbf{M}\{\exp(s^2)\} - \{\mathbf{M}(\hat{\theta}_1)\}^2 =$$

$$= \exp\left(2\mu + \frac{\sigma^2}{n}\right)\left[\exp\left\{\frac{\sigma^2}{n}\right\}\left(1 - \frac{2\sigma^2}{n}\right)^{-(n-1)/2} - \left(1 - \frac{\sigma^2}{n}\right)^{-(n-1)}\right],$$

что дает асимптотическую дисперсию

$$\mathbf{D}\hat{\theta}_1 \sim \exp(2\mu + \sigma^2) \cdot \frac{1}{n}\left(\sigma^2 + \frac{1}{2}\sigma^4\right).$$

Показать, что такую же асимптотическую дисперсию имеет оценка $\bar{\theta}_1$, введенная в упражнении 18.8. Вывести отсюда, что эффективность несмещенной моментной оценки для θ_1

$$\bar{y} = \frac{1}{n}\sum y$$

равна

$$\left(\sigma^2 + \frac{1}{2}\sigma^4\right)\Big/\{\exp(\sigma^2) - 1\}.$$

(Результаты упражнений 18.7—18.9 принадлежат Финни (1941) и Сичелу (1951, 1952).)

18.10 Пусть мультиноминальное распределение имеет n классов, причем все классы имеют равную вероятность $1/n$. Предположим, что в выборке объема N появилось k классов. Показать, что ФП для оценивания n имеет вид

$$L(k \mid n) = \frac{N!}{\prod_{i=1}^{n}(r_i!)}\left(\frac{1}{n}\right)^N \cdot \binom{n}{k} \frac{k!}{\prod_{j=1}^{N}(m_j!)},$$

где r_i $(\geqslant 0)$ — число наблюдений i-го класса и m_j — число классов, появившихся j $(\geqslant 1)$ раз. Показать далее, что МП-оценка $\hat{n}$ для n определяется соотношением

$$\frac{N}{\hat{n}} = \sum_{j=\hat{n}-k+1}^{\hat{n}} \frac{1}{j},$$

которое приближенно эквивалентно соотношению

$$\frac{N}{\hat{n}} = \log\left(\frac{\hat{n}}{\hat{n} - k + 1}\right),$$

и что статистика k достаточна для n. Показать, что для больших N

$$\mathbf{D}\hat{n} \sim \frac{n}{\exp\left(\frac{N}{n}\right) - \left(1 + \frac{N}{n}\right)}.$$

(Левонтин и Праут, 1956.)

18.11 Проверить, что в примере 18.14 МП-оценки (18.78) совместно достаточны для пяти параметров распределения, а МП-оценки (18.74), (18.75) совместно достаточны для σ_1^2, σ_2^2 и ρ.

18.12 Показать, что в случае оценивания параметра корреляции ρ двумерного нормального распределения при известных других четырех параметрах (примеры 18.14 и 18.15, случай (а)) выборочный коэффициент корреляции (который является МП-оценкой для ρ, если оцениваются все пять параметров, — см. (18.78)) имеет эффективность $1/(1+\rho^2)$. Показать далее, что если используется оценка

$$r' = \frac{\frac{1}{n}\sum (x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2},$$

то эффективность становится еще меньше и дается выражением $\left(\frac{1-\rho^2}{1+\rho^2}\right)^2$.

(Стьюарт, 1955а.)

18.13 Показать, что в упражнении 18.12 пара $\left(\sum x^2 + \sum y^2, \sum xy\right)$ является совместно достаточной для одного параметра ρ и что МП-оценка $\hat{\rho}$ есть функция от этой пары. Для $n = 1$ показать, что эта достаточная статистика является функцией от пары (x, y), которая сама достаточна, но что обратное неверно.

18.14 В примерах 18.14 и 18.15 найти МП-оценки для μ_1 при известных других параметрах и для σ_1^2 также при известных других параметрах. Показать, что асимптотические дисперсии этих оценок равны соответственно $\sigma_1^2(1-\rho^2)/n$ и $\frac{4\sigma_1^4(1-\rho^2)}{n(2-\rho^2)}$. Найти совместные МП-оценки для μ_1 и σ_1^2 при известных оставшихся трех параметрах и найти соответствующую им асимптотическую матрицу рассеяния.

18.15 В примере 18.17 вывести результаты (18.109) для некоррелированных оценок $\hat{\alpha}_u$ и $\hat{\beta}_u$, измеренных относительно центра расположения.

18.16 Показать, что центр расположения распределения Пирсона типа IV

$$dF \propto \exp\left\{-\nu \operatorname{arctg}\left(\frac{x-\alpha}{\beta}\right)\right\}\left\{1+\left(\frac{x-\alpha}{\beta}\right)^2\right\}^{-(\rho+2)/2} dx, \quad -\infty \leqslant x \leqslant \infty$$

(ν и ρ предполагаются известными), находится слева от моды распределения на расстоянии $\frac{\nu\beta}{\rho+4}$. Доказать также, что асимптотическая дисперсия МП-оценки $\hat{\alpha}$ дается выражением

$$\frac{1}{\beta^2}\frac{(\rho+1)(\rho+2)(\rho+4)}{(\rho+4)^2+\nu^2}$$

и что, следовательно, эффективность метода моментов при оценке расположения кривой равна

$$\frac{\rho^2(\rho-1)\{(\rho+4)^2+\nu^2\}}{(\rho+1)(\rho+2)(\rho+4)(\rho^2+\nu^2)}.$$

(Фишер, 1921а.)

18.17 Для выборки объема n из распределения

$$dF = dx, \quad \theta - \frac{1}{2} \leqslant x \leqslant \theta + \frac{1}{2},$$

показать, что пара $(x_{(1)}, x_{(n)})$ достаточна для θ, что одномерной достаточной статистики не существует и что ФП имеет максимум при любом значении из интервала $\left(x_{(n)} - \frac{1}{2},\ x_{(1)} + \frac{1}{2}\right)$. Показать также, что средняя точка этого интервала является несмещенной оценкой для θ.

18.18 Пусть из бесконечной совокупности, содержащей долю π элементов, обладающих данным признаком, элементы извлекаются до тех пор, пока не появится a элементов с этим признаком. Пусть n обозначает соответствующий объем выборки. Показать, что распределение n имеет вид

$$\binom{n-1}{a-1}\pi^a(1-\pi)^{n-a}, \qquad n = a,\ a+1,\ a+2, \ldots,$$

и что МП-оценкой для π является a/n. Показать также, что эта оценка смещена и имеет асимптотическую дисперсию

$$\pi^2(1-\pi)/a.$$

18.19 Рассмотреть задачу оценивания дисперсии θ_2 логнормального распределения, введенного в упражнениях 18.7, 18.8. Показать, что МП-оценка

$$\hat{\theta}_2 = \exp(2\bar{x} + s^2)\{\exp(s^2) - 1\}$$

смещена, а модифицированная оценка

$$\tilde{\theta}_2 = \exp(2\bar{x})\left\{t(2s^2) - t\left(\frac{n-2}{n-1}s^2\right)\right\}$$

несмещена.

18.20 Показать, что для оценки из упражнения 18.19 асимптотически имеем

$$\mathrm{D}\hat{\theta}_2 \sim \frac{2\sigma^2}{n}\exp(4\mu + 2\sigma^2)\,[2\{\exp(\sigma^2) - 1\}^2 + \sigma^2\{2\exp(\sigma^2) - 1\}^2]$$

и что, следовательно, эффективность несмещенной моментной оценки

$$s_y^2 = \frac{1}{n-1}\sum (y - \bar{y})^2$$

равна

$$\frac{2\sigma^2\,[2\{\exp(\sigma^2) - 1\}^2 + \sigma^2\{2\exp(\sigma^2) - 1\}^2]}{\{\exp(\sigma^2) - 1\}^2\{\exp(4\sigma^2) - 2\exp(3\sigma^2) + 3\exp(2\sigma^2) - 4\}}.$$

(Финни, 1941.)

18.21 Для случая нормального распределения с известной дисперсией σ^2 и неизвестным средним μ, принимающим только целочисленные значения, показать, что МП-оценка $\hat{\mu}$ есть целое число, ближайшее к выборочному среднему $\bar{x}$, и что выборочное распределение этой оценки дается выражением

$$f(\hat{\mu}) = \frac{1}{\sqrt{2\pi}}\int_{\left(\hat{\mu}-\mu-\frac{1}{2}\right)\sqrt{n}/\sigma}^{\left(\hat{\mu}-\mu+\frac{1}{2}\right)\sqrt{n}/\sigma} e^{-t^2/2}\,dt.$$

Вывести отсюда, что $\hat{\mu}$ является несмещенной и состоятельной оценкой для μ с асимптотической дисперсией

$$\mathrm{D}\hat{\mu} \sim \left(\frac{8\sigma^2}{n\pi}\right)^{1/2}\exp\left(-\frac{n}{8\sigma^2}\right),$$

экспоненциально убывающей при возрастании n.

(Хаммерсли, 1950.)

18.22 В предыдущем упражнении определим статистику T соотношением

$$\frac{1}{2}T = \text{целое число, ближайшее к } \frac{1}{2}\bar{x}.$$

Показать, что

$$\begin{aligned} \mathrm{D}T &< \mathrm{D}\hat{\mu}, && \text{если } \mu \text{ четно,} \\ \mathrm{D}T &\sim 1, && \text{если } \mu \text{ нечетно.} \end{aligned}$$

Вывести отсюда, что T состоятельна или несостоятельна в соответствии с тем, является μ четным числом или нечетным.

(Хаммерсли, 1950; дискуссия (Ч. Стейн).)

18.23 Пусть величина $\log(t-\gamma)$, $\gamma < t \leqslant \infty$, нормально распределена со средним μ и дисперсией σ^2 и оцениваются все три параметра (γ, μ, σ^2). Определяя $\hat{\mu}(\gamma)$ и $\hat{\sigma}^2(\gamma)$ соотношениями

$$\hat{\mu}(\gamma) = \frac{1}{n}\sum_{i=1}^{n} \log(t_i - \gamma) \quad \text{и} \quad \hat{\sigma}^2(\gamma) = \frac{1}{n}\sum_{i=1}^{n} \{\log(t_i - \gamma) - \hat{\mu}(\gamma)\}^2,$$

показать, что

$$L^{**}(\gamma) = \sup_{\mu,\,\sigma^2} L(x \mid \gamma, \mu, \sigma^2) \propto \{\hat{\sigma}(\gamma)\}^{-n} \prod_{i=1}^{n} (t_i - \gamma)^{-1}$$

и что если $t_{(1)}$ есть наименьшее наблюденное значение t, то

$$\lim_{\gamma \to t_{(1)}} L^{**}(\gamma) = +\infty,$$

так что МП-оценка для (γ, μ, σ^2) в случае рассматриваемого трехпараметрического логнормального распределения всегда равна $(t_{(1)}, -\infty, +\infty)$.

(Хилл, 1963а.)

18.24 Пусть имеется выборка объема n из распределения $f(x \mid \theta)$, сгруппированного в интервалы шириной h. Обозначим

$$f(x \mid \theta, h) = \int_{x-h/2}^{x+h/2} f(y \mid \theta)\, dy.$$

Используя разложение Тейлора, показать, что

$$f(x \mid \theta, h) = hf(x \mid \theta)\left\{1 + \frac{h^2}{24}\frac{\dfrac{\partial^2 f(x \mid \theta)}{\partial x^2}}{f(x \mid \theta)} + \ldots\right\}$$

и что, следовательно, с точностью до первого порядка поправка для МП-оценки, учитывающая группировку, равна

$$\Delta = -\frac{1}{24}h^2 \frac{\displaystyle\sum_{i=1}^{n} \frac{\partial}{\partial\theta}\left(\frac{\partial^2 f/\partial x^2}{f}\right)}{\displaystyle\sum_{i=1}^{n} \frac{\partial^2}{\partial\theta^2}(\log f)}$$

(значение правой части берется в точке $\hat{\theta}$).

(Линдли, 1950.)

18.25 Используя последнее упражнение, показать, что в случае оценивания среднего значения нормальной совокупности с известной дисперсией $\Delta = 0$, тогда как в случае оценивания дисперсии при известном среднем $\Delta = -h^2/12$.

Каждая из этих поправок есть в точности поправка Шеппарда на группировку для соответствующих генеральных моментов. Чтобы показать, что поправка на группировку для МП-оценок в общем случае не совпадает с поправкой Шеппарда, рассмотреть распределение

$$dF = e^{-x/\theta}\, dx/\theta, \qquad \theta > 0, \qquad 0 \leqslant x \leqslant \infty,$$

для которого $\hat{\theta} = \bar{x}$ (выборочное среднее). Поправка здесь равна

$$\Delta = -\frac{1}{12}\frac{h^2}{\bar{x}},$$

тогда как поправка Шеппарда для генерального среднего равна нулю.

(Линдли, 1950.)

18.26 Записывая отрицательное биномиальное распределение (5.33) в виде

$$f_r = \left(1 + \frac{m}{k}\right)^{-k} \binom{k+r-1}{r} \left(\frac{m}{m+k}\right)^r, \qquad r = 0, 1, 2, \ldots,$$

показать, что в случае выборки, состоящей из n независимых наблюдений, МП-оценка для m есть выборочное среднее

$$\hat{m} = \bar{r},$$

тогда как МП-оценка для k дается корнем уравнения

$$n \log\left(1 + \frac{\bar{r}}{k}\right) = \sum_{r=1}^{\infty} n_r \sum_{i=0}^{r-1} \frac{1}{k+i}$$

(n_r обозначает число наблюдений, принявших значение r; предполагается, что $n_0 < n$). Показать, что при уменьшающемся и стремящемся к нулю k правая часть уравнения превосходит левую и что если выборочная дисперсия s_r^2 превосходит $\bar{r}$, то левая часть превосходит правую при $k \to \infty$, т. е. уравнение имеет по крайней мере один конечный положительный корень. Показать, что, с другой стороны, если $s_r^2 < \bar{r}$, то обе части уравнения стремятся к равенству при $k \to \infty$, так что $\hat{k} = \infty$ и f_r сводится к распределению Пуассона с параметром m.

(Энскомб, 1950.)

18.27 Показать, что в предыдущем упражнении

$$\mathrm{D}\hat{m} = (m + m^2/k)/n,$$

$$\mathrm{D}\hat{k} \sim \left\{ \frac{2k(k+1)}{n\left(\frac{m}{m+k}\right)^2} \right\} \Bigg/ \left\{ 1 + 2\sum_{j=2}^{\infty} \frac{\left(\frac{j}{j+1}\right)\left(\frac{m}{m+k}\right)^{j-1}}{\binom{k+j}{j-1}} \right\},$$

$$\mathrm{cov}\,(\hat{m}, \hat{k}) \sim 0.$$

(Энскомб (1950); Фишер (1941) исследовал эффективность метода моментов в этом случае.)

18.28 Показать, что для распределения Неймана типа A, рассмотренного в упражнении 5.7, МП-оценки параметров λ_1 и λ_2 даются корнями системы уравнений

$$\hat{\lambda}_1 = \sum_r n_r (r+1) f_{r+1}/(n f_r),$$

$$\hat{\lambda}_1 \hat{\lambda}_2 = \bar{r},$$

где n_r, $\bar{r}$ имеют тот же смысл, что и в упражнении 18.26.

(Шентон (1949) в этом случае исследовал эффективность метода моментов, которая оказалась относительно низкой ($<70\%$) для малых λ_1 (<3) и больших λ_2 ($\geqslant 1$). Он изучил также вопрос об эффективности использования моментов в общем случае и в частном случае рядов Грама — Шарлье типа A (1950, 1951). См. также Кэтти и Гёрлэнд (1962).)

18.29 Показать, что МП-оценки параметров α и β «логистического» распределения

$$F(x) = \frac{1}{1+\exp\{-(\alpha+\beta x)\}}, \qquad -\infty \leqslant x \leqslant \infty,$$

даются корнями системы уравнений

$$\bar{x} = \frac{1}{\beta} + \sum_{i=1}^{n} \frac{x_i \exp(-\beta x_i)}{1+\exp\{-(\alpha+\beta x_i)\}} \Bigg/ \sum_{i=1}^{n} \frac{\exp(-\beta x_i)}{1+\exp\{-(\alpha+\beta x_i)\}},$$

$$e^{\alpha} = \frac{2}{n} \sum_{i=1}^{n} \frac{\exp(-\beta x_i)}{1+\exp\{-(\alpha+\beta x_i)\}}.$$

(Другие оценки для этого распределения были рассмотрены Гуптой и Гнанадесиканом (1966), использовавшими методы пунктов **19.18—20** (см. упражнение 32.12 ниже), и Блэмом (1958).)

18.30 Пусть имеются две независимые выборки объемов n_1 и n_2 из двух нормальных распределений с одинаковыми средними μ и дисперсиями $\lambda\sigma^2$ и σ^2. Найти МП-оценку для μ и показать, что ее асимптотическая дисперсия равна

$$\mathbf{D}(\hat{\mu}) = \sigma^2 \Big/ \left(\frac{n_1}{\lambda} + n_2\right).$$

Вывести отсюда, что несмещенная оценка

$$t = (n_1 \bar{x}_1 + n_2 \bar{x}_2)/(n_1 + n_2)$$

имеет эффективность

$$\frac{\lambda (n_1+n_2)^2}{(n_1\lambda + n_2)(n_1 + n_2\lambda)},$$

которая равна 1 тогда и только тогда, когда $\lambda = 1$

18.31 Показать, что для выборки объема n из нормального распределения со средним θ и дисперсией $V(\theta)$ МП-оценка $\hat{\theta}$ дается корнем уравнения

$$V' = 2(\bar{x} - \theta) + \frac{V'}{V}\frac{1}{n}\sum (x-\theta)^2.$$

Вывести отсюда, что если $V(\theta) = \sigma^2\theta^k$, где σ^2 известно, то $\hat{\theta}$ зависит как от $\bar{x}$, так и от $\sum x^2$, кроме случая $k = 0$ (когда $\hat{\theta}$ равно выборочному среднему $\bar{x}$, являющемуся одномерной достаточной статистикой) и случая $k = 1$ (когда $\hat{\theta}$ зависит только от $\sum x^2$).

18.32 Пусть в качестве оценки параметра θ берется θ^*, корень уравнения $g(x, \theta) = 0$. Показать, что если выполнены условия регулярности пунктов **17.14—15**, использованные при выводе выражения для МГД, и $\mathbf{M}\{g(x, \theta)\} = 0$, то справедливо неравенство

$$\frac{\mathbf{M}(g^2)}{\left\{\mathbf{M}\left(\frac{\partial g}{\partial\theta}\right)\right\}^2} \geqslant \frac{1}{\mathbf{M}\left\{\left(\frac{\partial \log L}{\partial\theta}\right)^2\right\}},$$

которое обращается в равенство только в случае, когда $g(x, \theta)$ отличается от $\partial \log L/\partial\theta$ постоянным множителем. В этом случае θ^* совпадает с МП-оценкой $\hat{\theta}$. Рассмотренное обобщение сводится к (17.22), когда $g(x, \theta) = t - \tau(\theta)$.

(Годэмб, 1960; Дёрбин, 1960.)

18.33 Показать, что для случайной выборки из распределения

$$f(x \mid \theta) = \begin{cases} (1/3)\exp\{-|x-\theta|\}, & x < \theta, \\ 1/3, & \theta \leqslant x \leqslant \theta + 1, \\ (1/3)\exp\{-|x-(\theta+1)|\}, & \theta + 1 < x, \end{cases}$$

МП-оценка для θ всегда единственна. (Рассмотреть, в частности, случаи $n = 1$, $n = 2$.)

(Даниэлс, 1961.)

18.34 Выборка, включающая n наблюдений, извлечена из нормальной совокупности со средним μ и дисперсией, которая с равными вероятностями может быть равна 1 или σ^2. Показать, что при $n \to \infty$ МП-оценки для (μ, σ^2) не существует.

(Кифер и Волфовиц, 1956)

18.35 Выборка объема n взята из распределения с непрерывной функцией плотности $f(x)$, определенной на интервале $0 \leqslant x \leqslant 1$ и принимающей значения, удовлетворяющие неравенству $0 \leqslant f(x) \leqslant 2$. Показать, что оценка $\hat{F}(x)$ для $F(x)$ является МП-оценкой тогда и только тогда, когда $\int_0^1 \hat{f}(x)\,dx = 1$, $\hat{f}(x)$ непрерывна и $\hat{f}(x_i) = 2$, $i = 1, 2, \ldots, n$. Вывести отсюда, что существует много МП-оценок как состоятельных, так и несостоятельных.

(Бахадур, 1958.)

ГЛАВА 19

ОЦЕНИВАНИЕ: МЕТОД НАИМЕНЬШИХ КВАДРАТОВ И ДРУГИЕ МЕТОДЫ

19.1 В этой главе будут рассмотрены принципы оценивания, отличные от принципа максимума правдоподобия (МП), которому была посвящена глава 18. Наиболее важным из них является принцип (или метод) наименьших квадратов (сокращенно: НК). Этот метод, отличаясь по подходу от метода МП и обладая своими собственными оптимальными свойствами, в то же время совпадает с методом МП в важном случае нормально распределенных наблюдений. Другие методы, рассматриваемые в этой главе, существенно отличаются от метода МП и если оказываются эквивалентными ему, то только в асимптотическом смысле.

Метод наименьших квадратов

19.2 Мы знаем (примеры 18.2, 17.6), что МП-оценка для среднего μ по выборке объема n из нормального распределения

$$dF(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right\} dy \qquad (19.1)$$

получается максимизацией по μ функции правдоподобия

$$\log L(y \mid \mu) = -\frac{1}{2} n \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{j=1}^{n} (y_j - \mu)^2. \qquad (19.2)$$

Из (19.2) видно, что эта функция максимальна, когда минимальна сумма

$$\sum_{j=1}^{n} (y_j - \mu)^2. \qquad (19.3)$$

Таким образом, принцип МП сводится к тому, чтобы выбирать такое $\hat{\mu}$, которое обращало бы в минимум (19.3).

Предположим теперь, что генеральное среднее μ является линейной функцией параметров θ_i $(i = 1, 2, \ldots, k)$:

$$\mu = \sum_{i=1}^{k} x_i \theta_i, \qquad (19.4)$$

где x_i обозначают не случайные величины, а известные постоянные коэффициенты при неизвестных параметрах θ_i. Для того чтобы в этом случае оценить каждый из параметров θ_i, надо согласно (19.3) и (19.4) минимизировать по θ_i сумму

$$\sum_{j=1}^{n}\left(y_j-\sum_{i=1}^{k}x_i\theta_i\right)^2. \tag{19.5}$$

Можно пойти еще дальше по пути обобщения. Предположим, что n наблюдений получены не из одинаковых нормальных распределений, а из распределений с разными средними. Именно, пусть

$$\mu_j=\sum_{i=1}^{k}x_{ij}\theta_i, \quad j=1, 2, \ldots, n. \tag{19.6}$$

В этом случае надо минимизировать по θ_i сумму

$$\sum\left(y_j-\sum_i x_{ij}\theta_i\right)^2. \tag{19.7}$$

19.3 Метод НК получил свое название из-за его связи с минимизацией сумм квадратов вида (19.7). Как общий принцип он формулируется следующим образом: в качестве оценки вектора θ в некотором выражении $p(x, \theta)=0$, где x обозначает наблюдение, следует взять такое $\hat{\theta}$, которое обращает в минимум сумму

$$\sum_{j=1}^{n}\{p(x_j, \hat{\theta})\}^2.$$

Как и в случае любого другого систематического принципа оценивания, вопрос о возможности использования метода НК зависит от свойств оценок, к которым он приводит. В противоположность методу МП, этот метод в общем случае не обладает даже асимптотическими оптимальными свойствами. Однако в одном очень важном классе ситуаций он даже при малых выборках обладает свойством оптимальности, состоящим в том, что он дает несмещенные оценки, являющиеся линейными функциями от наблюдений и имеющие минимальную дисперсию (МД). Это случай так называемой линейной модели, когда наблюдения имеют одинаковые дисперсии и (возможно, разные) средние, являющиеся линейными функциями неизвестных параметров, и когда наблюдения попарно некоррелированы. Именно таковы были предположения в модели (19.6). Однако теперь мы откажемся от предположения нормальности, на котором были основаны рассуждения в **19.2**, поскольку оно совершенно не нужно для установления оптимальных свойств оценок, полу-

ченных по методу НК (НК-оценок). Перейдем к точной постановке задачи. При этом окажется удобным, как и в главе 15, использовать обозначения и терминологию матричной и векторной теории.

Оценки по методу наименьших квадратов для линейной модели

19.4 Мы запишем линейную модель в виде

$$\boldsymbol{y} = \boldsymbol{X}\theta + \boldsymbol{\varepsilon}, \tag{19.8}$$

где $\boldsymbol{y}$ — вектор-столбец наблюдений размерности n, $\boldsymbol{X}$ — матрица $(n \times k)$ известных коэффициентов $(n > k)$, θ — вектор-столбец параметров размерности k и $\boldsymbol{\varepsilon}$ — вектор-столбец случайных «ошибок» размерности n с математическим ожиданием

$$\mathsf{M}(\boldsymbol{\varepsilon}) = 0 \tag{19.9}$$

и матрицей рассеяния

$$\boldsymbol{V}(\boldsymbol{\varepsilon}) = \mathsf{M}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma^2 \boldsymbol{I} \tag{19.10}$$

($\boldsymbol{I}$ — единичная матрица порядка $(n \times n)$). Условия (19.9) и (19.10) соответствуют предположению о том, что ε_i некоррелированы, имеют нулевые средние и одну и ту же дисперсию σ^2. Эти условия играют существенную роль при получении следующих далее результатов. Линейную модель можно обобщить на менее ограничительную ситуацию (см. **19.17** и упражнения 19.2 и 19.5), но результаты при этом соответственно изменятся.

Следует отметить, что на элементы $\boldsymbol{X}$ не наложено никаких ограничений. Поэтому, как мы увидим дальше в **33.43** и главе 35 (том 3), определяя соответствующим образом эти элементы, можно использовать рассматриваемую линейную модель для анализа категоризованных и классифицированных наблюдений.

Метод НК состоит в минимизации скалярной суммы квадратов

$$S = (\boldsymbol{y} - \boldsymbol{X}\theta)'(\boldsymbol{y} - \boldsymbol{X}\theta) \tag{19.11}$$

по компонентам вектора θ. Необходимым условием обращения (19.11) в минимум является условие $\partial S/\partial\theta = 0$. Выполняя дифференцирование, получаем $2\boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\theta) = 0$, откуда находим вектор НК-оценок:

$$\hat{\theta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}. \tag{19.12}$$

Предполагается, что матрица $(\boldsymbol{X}'\boldsymbol{X})$ невырождена и, следовательно, может быть обращена.

Пример 19.1

Рассмотрим простейший случай, когда θ имеет только одну компоненту. Тогда можно переписать (19.8) в виде

$$\boldsymbol{y} = \boldsymbol{x}\theta + \boldsymbol{\varepsilon},$$

где $\boldsymbol{x}$ — вектор-столбец размерности n. НК-оценка согласно (19.12) дается формулой

$$\hat{\theta} = (\boldsymbol{x}'\boldsymbol{x})^{-1}\boldsymbol{x}'\boldsymbol{y} = \sum_{j=1}^{n} x_j y_j \Big/ \sum_{j=1}^{n} x_j^2.$$

Пример 19.2

Предположим, что θ имеет две компоненты. Матрица $\boldsymbol{X}$ состоит в этом случае из двух векторов $\boldsymbol{x}_1$ и $\boldsymbol{x}_2$, и модель (19.8) принимает вид

$$\boldsymbol{y} = (\boldsymbol{x}_1, \boldsymbol{x}_2)\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix} + \boldsymbol{\varepsilon}.$$

Пользуясь далее формулой (19.12), получаем следующее выражение для НК-оценки:

$$\hat{\theta} = \begin{pmatrix}\boldsymbol{x}_1'\boldsymbol{x}_1 & \boldsymbol{x}_1'\boldsymbol{x}_2\\ \boldsymbol{x}_2'\boldsymbol{x}_1 & \boldsymbol{x}_2'\boldsymbol{x}_2\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{x}_1'\boldsymbol{y}\\ \boldsymbol{x}_2'\boldsymbol{y}\end{pmatrix} = \begin{pmatrix}\sum x_1^2 & \sum x_1x_2\\ \sum x_1x_2 & \sum x_2^2\end{pmatrix}^{-1}\begin{pmatrix}\sum x_1y\\ \sum x_2y\end{pmatrix}$$

(все суммирования ведутся по $j = 1, 2, \ldots, n$). Замечая, что элементами матрицы $(\boldsymbol{X}'\boldsymbol{X})$ являются скалярные произведения векторов-столбцов матрицы $\boldsymbol{X}$ друг на друга и самих на себя и что элементами вектора $\boldsymbol{X}'\boldsymbol{y}$ являются скалярные произведения вектора $\boldsymbol{y}$ на каждый из векторов $\boldsymbol{x}$, легко обобщить этот пример на случай произвольного числа компонент.

Пример 19.3

Пусть в примере 19.2 $\boldsymbol{x}_1 = \mathbf{1}$, где $\mathbf{1}$ обозначает вектор, все компоненты которого равны единице. Тогда

$$\hat{\theta} = \begin{pmatrix}n & \sum x_2\\ \sum x_2 & \sum x_2^2\end{pmatrix}^{-1}\begin{pmatrix}\sum y\\ \sum x_2y\end{pmatrix}.$$

Обращение первой матрицы дает

$$\begin{pmatrix}n & \sum x_2\\ \sum x_2 & \sum x_2^2\end{pmatrix}^{-1} = \frac{1}{\{n\sum x_2^2 - (\sum x_2)^2\}}\begin{pmatrix}\sum x_2^2 & -\sum x_2\\ -\sum x_2 & n\end{pmatrix}.$$

Умножая полученное выражение на $\begin{pmatrix} \sum y \\ \sum x_2 y \end{pmatrix}$, получаем

$$\hat{\theta} = \frac{1}{\{n \sum x_2^2 - (\sum x_2)^2\}} \begin{pmatrix} \sum x_2^2 \sum y - \sum x_2 \sum x_2 y \\ -\sum x_2 \sum y + n \sum x_2 y \end{pmatrix},$$

или

$$\hat{\theta}_1 = \frac{\sum x_2^2 \sum y - \sum x_2 \sum x_2 y}{n \sum x_2^2 - (\sum x_2)^2}, \qquad \hat{\theta}_2 = \frac{n \sum x_2 y - \sum x_2 \sum y}{n \sum x_2^2 - (\sum x_2)^2}.$$

Упрощая, имеем

$$\hat{\theta}_2 = \frac{\frac{1}{n} \sum x_2 y - \frac{1}{n} \sum x_2 \cdot \frac{1}{n} \sum y}{\frac{1}{n} \sum x_2^2 - \left(\frac{1}{n} \sum x_2\right)^2} = \frac{\sum (x_2 - \bar{x}_2)(y - \bar{y})}{\sum (x_2 - \bar{x}_2)^2},$$

откуда

$$\hat{\theta}_1 = \bar{y} - \hat{\theta}_2 \bar{x}_2.$$

Мы видим, что $\hat{\theta}_2$ имеет в точности такой же вид, как $\hat{\theta}$ в примере 19.1, с той лишь разницей, что отклонения от начала координат заменяются отклонениями от средних. Вообще такой эффект наблюдается всегда, когда вводится новый параметр, которому соответствует вектор $\boldsymbol{x}$, состоящий из единиц (см. упражнение 19.1).

19.5 Мы можем теперь доказать несмещенность НК-оценок (19.12). Используя (19.8), можно переписать (19.12) в виде

$$\hat{\theta} = (\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{X}'(\boldsymbol{X}\theta + \varepsilon) = \theta + (\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{X}'\varepsilon. \qquad (19.13)$$

Поскольку матрица $\boldsymbol{X}$ постоянна, мы получаем в силу (19.9) требуемый результат:

$$\mathsf{M}(\hat{\theta}) = \theta. \qquad (19.14)$$

Матрица рассеяния $\hat{\theta}$

$$\boldsymbol{V}(\hat{\theta}) = \mathsf{M}\{(\hat{\theta} - \theta)(\hat{\theta} - \theta)'\}$$

после подстановки в нее (19.13) принимает вид

$$\boldsymbol{V}(\hat{\theta}) = \mathsf{M}\{[(\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{X}'\varepsilon][(\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{X}'\varepsilon]'\} =$$
$$= (\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{X}' \mathsf{M}(\varepsilon\varepsilon') \boldsymbol{X} (\boldsymbol{X}'\boldsymbol{X})^{-1}, \qquad (19.15)$$

откуда, учитывая (19.10), получаем

$$\boldsymbol{V}(\hat{\theta}) = \sigma^2 (\boldsymbol{X}'\boldsymbol{X})^{-1}. \qquad (19.16)$$

(19.12) и (19.16) показывают, что как вычисление НК-оценки, так и вычисление ее матрицы рассеяния существенно зависят

от обращения матрицы $(X'X)$. В простых случаях (см. пример 19.3) обратная матрица может быть получена алгебраическим путем. Однако в более сложных случаях приходится применять численные методы, подобные описанным Л. Фоксом (1950) и Л. Фоксом и Хейесом (1951).

Пример 19.4

Дисперсия $\hat{\theta}$ в примере 19.1 согласно (19.16) равна

$$\mathsf{D}(\hat{\theta}) = \sigma^2 / \sum x_j^2.$$

Пример 19.5

В примере 19.2 имеем

$$(X'X)^{-1} = \begin{pmatrix} \sum x_1^2 & \sum x_1 x_2 \\ \sum x_1 x_2 & \sum x_2^2 \end{pmatrix}^{-1} =$$

$$= \frac{1}{\left\{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2\right\}} \begin{pmatrix} \sum x_2^2 & -\sum x_1 x_2 \\ -\sum x_1 x_2 & \sum x_1^2 \end{pmatrix},$$

откуда согласно (19.16)

$$\mathsf{D}(\hat{\theta}_1) = \frac{\sigma^2 \sum x_2^2}{\left\{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2\right\}} = \frac{\sigma^2}{\sum x_1^2 \left\{1 - \dfrac{\left(\sum x_1 x_2\right)^2}{\sum x_1^2 \sum x_2^2}\right\}}.$$

Для получения $\mathsf{D}(\hat{\theta}_2)$ надо в последнем выражении поменять местами индексы 1 и 2. Ковариация равна

$$\operatorname{cov}(\hat{\theta}_1, \hat{\theta}_2) = \frac{-\sigma^2 \sum x_1 x_2}{\left\{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2\right\}};$$

она обращается в нуль в том и только в том случае, когда $\sum x_1 x_2 = 0$.

Пример 19.6

В примере 19.3 мы нашли, что

$$(X'X)^{-1} = \frac{1}{\sum (x_2 - \bar{x}_2)^2} \begin{pmatrix} \frac{1}{n}\sum x_2^2 & -\bar{x}_2 \\ -\bar{x}_2 & 1 \end{pmatrix},$$

откуда согласно (19.16)

$$\mathsf{D}(\hat{\theta}_1) = \sigma^2 \sum x_2^2 / \left\{n \sum (x_2 - \bar{x}_2)^2\right\},$$

$$\mathsf{D}(\hat{\theta}_2) = \sigma^2 / \sum (x_2 - \bar{x}_2)^2,$$

$$\operatorname{cov}(\hat{\theta}_1, \hat{\theta}_2) = -\sigma^2 \bar{x}_2 / \sum (x_2 - \bar{x}_2)^2.$$

Как и следовало ожидать, $\mathsf{D}(\hat{\theta}_2)$ равна $\mathsf{D}(\hat{\theta})$ в примере 19.4 с заменой отклонений от начала отсчета на отклонения от среднего в знаменателе. $\hat{\theta}_1$ и $\hat{\theta}_2$ некоррелированы в том и только в том случае, когда $\bar{x}_2 = 0$.

Свойства оптимальности

19.6 Теперь мы покажем, что несмещенные линейные МД-оценки для любой системы линейных функций от параметров θ_i даются методом НК *). Красивое доказательство этого факта имеется в работе Плэкетта (1949), посвященной истокам теории НК и вскрывающей основополагающую роль результатов Гаусса в этой области.

Пусть t — любой вектор оценок, линейных относительно наблюдений y, т. е. имеющий вид

$$t = Ty. \tag{19.17}$$

Мы требуем, чтобы оценка t была несмещенной для некоторой системы линейных функций от параметров, скажем $C\theta$, т. е. $\mathsf{M}(t) = \mathsf{M}(Ty) = C\theta$ (C — известная матрица коэффициентов), что ввиду (19.8) дает

$$\mathsf{M}\{T(X\theta + \varepsilon)\} = C\theta. \tag{19.18}$$

Отсюда, используя (19.9) и учитывая, что (19.18) должно выполняться при всех θ, получаем

$$TX = C. \tag{19.19}$$

(9.19) является необходимым и достаточным условием того, чтобы оценка Ty была несмещенной для $C\theta$.

Поскольку ввиду (19.17), (19.8) и (19.19) $t - C\theta = T\varepsilon$, то матрицу рассеяния

$$V(t) = \mathsf{M}\{(t - C\theta)(t - C\theta)'\}, \tag{19.20}$$

учитывая (19.10), можно записать в виде

$$V(t) = \mathsf{M}(T\varepsilon\varepsilon'T') = \sigma^2 TT'. \tag{19.21}$$

Нам надо минимизировать диагональные элементы в (19.21), являющиеся дисперсиями нашей системы оценок.

Имеет место тождество

$$TT' \equiv \{C(X'X)^{-1}X'\}\{C(X'X)^{-1}X'\}' + \\ + \{T - C(X'X)^{-1}X'\}\{T - C(X'X)^{-1}X'\}', \tag{19.22}$$

которое легко проверить, производя умножение в правой части, являющейся суммой двух слагаемых вида AA'. Каждое из этих

*) Барнард (1963) приводит и другие свойства оптимальности метода НК.

двух слагаемых имеет, следовательно, неотрицательные диагональные элементы. Однако только второе слагаемое в правой части (19.22) зависит от $\boldsymbol{T}$. Поэтому диагональные элементы суммы будут минимальными, когда диагональные элементы второго слагаемого будут равны нулю. Это будет иметь место, если выполнено соотношение

$$\boldsymbol{T} = \boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'. \tag{19.23}$$

Таким образом, в силу (19.17) и (19.23) несмещенный вектор МД-оценок для $\boldsymbol{C}\theta$ дается выражением

$$\boldsymbol{t} = \boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}, \tag{19.24}$$

которое получается просто заменой θ на его НК-оценку (19.12), т. е.

$$\boldsymbol{t} = \boldsymbol{C}\hat{\theta}. \tag{19.25}$$

В силу (19.21) и (19.23) матрица рассеяния этой оценки равна

$$\boldsymbol{V}(\boldsymbol{t}) = \sigma^2\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{C}'. \tag{19.26}$$

Если $\boldsymbol{C} = \boldsymbol{I}$ (единичная матрица), т. е. оценивается сам вектор θ, то (19.24) и (19.26) сводятся соответственно к (19.12) и (19.16).

НК-оценки являются несмещенными МД-оценками в линейной модели в классе линейных функций от $\boldsymbol{y}$, но могут не обладать этим свойством, если в качестве оценок допускаются также и нелинейные функции от $\boldsymbol{y}$. Это зависит от распределения ε_i. Если последние распределены нормально (и, следовательно, независимы), то НК-оценки сохраняют свое свойство и в классе всех оценок, поскольку они являются функциями от минимальной системы из $k+1$ достаточных статистик $(\hat{\theta}, s^2)$ для $k+1$ параметра (см. **17.39**). Андерсон (1962) показал, что если рассматриваются все возможные распределения ε_i, то НК-оценки очень редко являются МД-оценками в классе всех оценок.

19.7 Поучительно, следуя Дёрбину и Кендаллу (1951), рассмотреть геометрический подход к свойству МД. Здесь будет рассмотрен лишь простейший случай, когда оценивается один параметр θ по выборке из n наблюдений с одинаковыми средними θ и дисперсиями σ^2. В этом случае $y_j = \theta + \varepsilon_j$, $j = 1, 2, \ldots, n$, что представляет собой равенство (19.8), в котором $k = 1$, а $\boldsymbol{X}$ — вектор-столбец размерности n, состоящий из единиц. Рассматриваются линейные оценки вида

$$t = \sum_j c_j y_j \tag{19.27}$$

(простейший частный случай формы (19.17)). Условие несмещенности (19.19) здесь имеет вид

$$\sum_j c_j = 1. \tag{19.28}$$

Рассмотрим n-мерное евклидово пространство, в котором каждому c_j соответствует одна координата. Мы назовем это пространство пространством оценок. (19.28) является уравнением гиперплоскости в этом пространстве. Каждая точка P этой гиперплоскости однозначно соответствует некоторой несмещенной оценке. Поскольку y_j некоррелированы, из (19.27) получаем

$$\mathrm{D}t = \sigma^2 \sum_j c_j^2, \tag{19.29}$$

т. е. дисперсия t равна $\sigma^2 OP^2$, где O — начало координат пространства оценок. Отсюда сразу следует, что t имеет МД, когда P есть основание перпендикуляра к гиперплоскости, проведенного из O. Из соображений симметрии тогда все c_j равны $1/n$ и t есть выборочное среднее $\bar{y}$.

Рассмотрим теперь обычное n-мерное выборочное пространство, в котором каждому y_j соответствует одна координата. Билинейная форма (19.27) устанавливает двойственность между этим пространством и пространством оценок. При любом фиксированном t точке в одном пространстве соответствует гиперплоскость в другом, тогда как при переменном t точке в одном пространстве соответствует семейство параллельных гиперплоскостей в другом. Гиперплоскости (19.28) в пространстве оценок отвечает точка $(t, t, \ldots, t)$ в выборочном пространстве, лежащая на векторе с равными компонентами. Если выходящий из начала вектор ортогонален гиперплоскости в одном пространстве, то соответствующие гиперплоскости и вектор ортогональны в другом пространстве.

Из сказанного вытекает, что несмещенная МД-оценка дается в выборочном пространстве гиперплоскостью, ортогональной к вектору с равными компонентами в точке $L = (\bar{y}, \bar{y}, \ldots, \bar{y})$. Если Q — выборочная точка, то мы опускаем перпендикуляр из Q на вектор с равными компонентами, чтобы найти L, т. е. минимизируем $QL^2 = \sum_j (y_j - t)^2$. Тем самым мы минимизируем сумму квадратов в выборочном пространстве и, следовательно, минимизируем дисперсию (другую сумму квадратов) в пространстве оценок в силу двойственности, установленной между ними.

Крускал (1961) развивает полностью геометрический подход к теории НК.

19.8 Из результата, полученного в **19.6**, сразу следует, что НК-оценка $\hat{\theta}$ минимизирует значение обобщенной дисперсии

линейных оценок для θ. Это утверждение, принадлежащее Эйткину (1948), справедливо при любом объеме выборки, в отличие от эквивалентного асимптотического результата, доказанного в **18.28** для МП-оценок. Мы приводим здесь доказательство этого факта, принадлежащее Даниэлсу (1951, 1952).

Результат пункта **19.6** в случае оценивания одной линейной функции $c'\theta$, где c' — вектор-строка размерности $1 \times k$, состоит в том, что

$$\mathsf{D}(c'\hat{\theta}) \leqslant \mathsf{D}(c'u), \tag{19.30}$$

где $\hat{\theta}$ — НК-оценка, а u — любая другая линейная оценка для θ. Можно переписать (19.30) в виде

$$c'Vc \leqslant c'Dc, \tag{19.31}$$

где V обозначает матрицу рассеяния $\hat{\theta}$, а D — матрицу рассеяния u.

Совершая действительное невырожденное преобразование

$$c = Ab, \tag{19.32}$$

одновременно приводящее V и D к диагональному виду, можно переписать (19.31) в виде

$$b'(A'VA)b \leqslant b'(A'DA)b, \tag{19.33}$$

где матрицы в скобках диагональны. Выбирая затем соответствующим образом b, легко заметить, что ни один из элементов $(A'VA)$ не может быть больше, чем соответствующий элемент $(A'DA)$. Таким образом,

$$|A'VA| \leqslant |A'DA|,$$

т. е.

$$|A'||V||A| \leqslant |A'||D||A|,$$

или

$$|V| \leqslant |D|, \tag{19.34}$$

что и требовалось доказать.

Оценивание дисперсии

19.9 Результат пункта **19.6** является первой частью утверждения, обычно называемого теоремой Гаусса о наименьших квадратах. Вторая часть теоремы относится к оцениванию дисперсии по наблюдениям.

Рассмотрим «остатки»

$$y - X\theta = [X\theta + \varepsilon] - X[(X'X)^{-1}X'(X\theta + \varepsilon)], \tag{19.35}$$

полученные при оценивании по методу НК (19.8) и (19.12). Члены с θ взаимно уничтожаются, и мы получаем

$$y - X\theta = \{I_n - X(X'X)^{-1}X'\}\varepsilon, \tag{19.36}$$

где I_n — единичная матрица порядка n.

Матрица в фигурных скобках в правой части (19.36) симметрична и идемпотентна, что легко проверить транспонированием и возведением ее в квадрат. Сумма квадратов остатков равна

$$(y - X\theta)'(y - X\theta) = \varepsilon'\{I_n - X(X'X)^{-1}X'\}\varepsilon. \tag{19.37}$$

Далее,

$$\varepsilon'B\varepsilon = \sum_i b_{ii}\varepsilon_i^2 + \sum_{i \neq j} b_{ij}\varepsilon_i\varepsilon_j.$$

Таким образом, ввиду (19.10)

$$\mathbf{M}(\varepsilon'B\varepsilon) = \sigma^2 \operatorname{tr} B. \tag{19.38}$$

Используя (19.38), мы получаем из (19.37)

$$\mathbf{M}\{(y - X\hat{\theta})'(y - X\hat{\theta})\} = \sigma^2 \operatorname{tr}\{I_n - X(X'X)^{-1}X'\} = \\ = \sigma^2[\operatorname{tr} I_n - \operatorname{tr}\{X(X'X)^{-1}X'\}]. \tag{19.39}$$

Меняя порядок матриц $X(X'X)^{-1}$ и X' под знаком оператора следа и превращая тем самым произведение из матрицы $(n \times n)$ в матрицу $(k \times k)$, мы можем записать правую часть (19.39) в виде

$$\sigma^2\{\operatorname{tr} I_n - \operatorname{tr} X' \cdot X(X'X)^{-1}\} = \sigma^2\{\operatorname{tr} I_n - \operatorname{tr} I_k\},$$

откуда

$$\mathbf{M}\{(y - X\hat{\theta})'(y - X\hat{\theta})\} = \sigma^2(n - k). \tag{19.40}$$

Таким образом, несмещенная оценка для σ^2 согласно (19.40) дается формулой

$$\frac{1}{n-k}(y - X\hat{\theta})'(y - X\hat{\theta}) = s^2, \tag{19.41}$$

т. е. равна сумме квадратов остатков, деленной на число наблюдений минус число оцениваемых параметров.

Полученный результат позволяет строить несмещенные оценки для матриц рассеяния (19.16) и (19.26), подставляя в эти выражения вместо σ^2 статистику s^2, определенную формулой (19.41).

Пример 19.7

Несмещенная оценка для σ^2 в примерах 19.1 и 19.4 согласно (19.41) имеет вид

$$s^2 = \frac{1}{n-1}\sum_{j=1}^{n}(y_j - x_j\hat{\theta})^2.$$

Следовательно, несмещенной оценкой для $\mathsf{D}(\hat{\theta})$ будет $s^2/\sum x_j^2$.

Пример 19.8

Несмещенная оценка для σ^2 в примерах 19.2 и 19.5 согласно (19.41) равна

$$s^2 = \frac{1}{n-2}\sum_{j=1}^{n}(y_j - x_{1j}\hat{\theta}_1 - x_{2j}\hat{\theta}_2)^2,$$

где

$$\begin{pmatrix}\hat{\theta}_1\\ \hat{\theta}_2\end{pmatrix} = \frac{1}{\{\sum x_1^2 \sum x_2^2 - (\sum x_1x_2)^2\}}\begin{pmatrix}\sum x_2^2 & -\sum x_1x_2\\ -\sum x_1x_2 & \sum x_1^2\end{pmatrix}\begin{pmatrix}\sum x_1y\\ \sum x_2y\end{pmatrix} =$$

$$= \frac{1}{\{\sum x_1^2 \sum x_2^2 - (\sum x_1x_2)^2\}}\begin{pmatrix}\sum x_2^2 \sum x_1y - \sum x_1x_2 \sum x_2y\\ \sum x_1^2 \sum x_2y - \sum x_1x_2 \sum x_1y\end{pmatrix},$$

что сводится к случаю, рассмотренному в примерах 19.3 и 19.6, если положить $x_{1j} = 1$.

Предположение нормальности

19.10 До сих пор мы не делали никаких предположений относительно распределений ошибок ε_i, кроме условий (19.9) и (19.10), касающихся первых и вторых моментов. Замечательно, что нам не понадобилось никаких предположений о виде распределений ошибок: мы получили несмещенные оценки параметров и несмещенные оценки выборочных дисперсий и ковариаций этих оценок без предположений о распределениях ошибок. Однако при проверке гипотез относительно параметров такие предположения необходимы. Вопросы проверки гипотез в линейной модели будут рассмотрены в главе 24. Сейчас мы подчеркнем лишь наиболее важные аспекты возникающей здесь ситуации.

19.11 Если предположить, что ε_i распределены нормально, то из их некоррелированности будет следовать их *независимость* (см. **15.3**), и тогда можно будет использовать утверждение пункта **15.11** о том, что идемпотентная квадратичная форма от независимых нормированных нормальных вариант имеет χ^2-распределение с числом степеней свободы, равным рангу

квадратичной формы. Применяя этот результат к сумме квадратов остатков (19.37), получим (в обозначениях (19.41)), что $(n-k)s^2/\sigma^2$ имеет χ^2-распределение с $n-k$ степенями свободы.

Далее, имеем тождество

$$\boldsymbol{y}'\boldsymbol{y} \equiv (\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\theta}})'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\theta}}) + (\boldsymbol{X}\hat{\boldsymbol{\theta}})'(\boldsymbol{X}\hat{\boldsymbol{\theta}}), \tag{19.42}$$

которое легко проверить, используя (19.12). Второе слагаемое в правой части (19.42) можно представить в виде

$$\hat{\boldsymbol{\theta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\theta}} = \boldsymbol{y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = (\boldsymbol{\varepsilon}' + \boldsymbol{\theta}'\boldsymbol{X}')\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon}). \tag{19.43}$$

Из (19.43) следует, что если $\boldsymbol{\theta} = \boldsymbol{0}$, то

$$\hat{\boldsymbol{\theta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\theta}} = \boldsymbol{\varepsilon}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}, \tag{19.44}$$

и тогда (19.42) можно переписать, используя (19.37) и (19.44), следующим образом:

$$\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}'\{\boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\}\boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}'\{\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\}\boldsymbol{\varepsilon}. \tag{19.45}$$

В **19.9** было показано, что ранг первой матрицы в фигурных скобках в правой части (19.45) равен $n-k$; там же было установлено, что ранг второй матрицы в фигурных скобках в (19.45) равен k. Таким образом, ранги в правой части дают в сумме n, ранг левой части, и, следовательно, применима теорема Кокрэна (15.16), согласно которой две квадратичные формы в правой части (19.54) независимы и имеют распределения (после деления на σ^2 в каждом случае) χ^2 с $n-k$ и k степенями свободы.

19.12 Отметим, что величина $(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\theta}})'(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\theta}})$ в **19.11** распределена по закону χ^2 при любом истинном значении $\boldsymbol{\theta}$, тогда как второе слагаемое $(\boldsymbol{X}\hat{\boldsymbol{\theta}})'(\boldsymbol{X}\hat{\boldsymbol{\theta}})$ в (19.42) имеет распределение χ^2 только при $\boldsymbol{\theta} = \boldsymbol{0}$. В любом случае из (19.43), учитывая (19.9), получаем

$$\mathsf{M}\{(\boldsymbol{X}\hat{\boldsymbol{\theta}})'(\boldsymbol{X}\hat{\boldsymbol{\theta}})\} = \mathsf{M}\{\boldsymbol{\varepsilon}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}\} + \boldsymbol{\theta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta}. \tag{19.46}$$

В **19.9** было показано, что первое слагаемое в правой части равно $k\sigma^2$. Следовательно,

$$\mathsf{M}\{(\boldsymbol{X}\hat{\boldsymbol{\theta}})'(\boldsymbol{X}\hat{\boldsymbol{\theta}})\} = k\sigma^2 + (\boldsymbol{X}\boldsymbol{\theta})'(\boldsymbol{X}\boldsymbol{\theta}), \tag{19.47}$$

что превосходит $k\sigma^2$, если только не выполнено равенство $\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{0}$, которое возможно лишь при $\boldsymbol{\theta} = \boldsymbol{0}$, за исключением некоторых специальных значений $\boldsymbol{X}$. Таким образом, интуитивно

разумно использовать отношение $(X\hat{\theta})'(X\hat{\theta})/(ks^2)$ (где статистика s^2, определенная формулой (19.41), всегда имеет математическое ожидание σ^2) для проверки гипотезы $\theta = 0$. Мы вернемся к более строгому обоснованию этой и подобных процедур в главе 24.

Случай вырождения

19.13 В **19.4** мы предполагали, что $X'X$ невырождена, так что (19.12) имело смысл. Предполагалось также, что $n > k$, так что имело смысл (19.41). Если $n = k$, то (19.12) остается в силе, если существует $(X'X)^{-1}$, но формула (19.41) становится бесполезной, поскольку сумма квадратов остатков, как показывает (19.37), тождественно обращается в нуль. Если $n < k$, то ранг X (и ранг $X'X$, равный рангу X) меньше k, так что матрица $X'X$ не имеет обратной.

Пусть теперь X имеет ранг $r < k$, и пусть $n \geqslant k$. В этом случае задачу оценивания по методу НК приходится решать заново, так как $X'X$ не имеет обратной матрицы, и поэтому (19.12) недействительно. Далее мы следуем Плэкетту (1950).

Условие (19.19) по-прежнему остается необходимым и достаточным условием того, чтобы оценка Ty была несмещенной для $C\theta$. В вырожденном случае оно, однако, не может быть выполнено, если мы хотим оценить сам вектор θ, поскольку тогда (19.19) принимает вид

$$TX = I. \tag{19.48}$$

Действительно, вспоминая, что матрица X имеет ранг r, произведем разбиение ее на блоки:

$$X = \begin{pmatrix} X_{r,r} & X_{r,k-r} \\ X_{n-r,r} & X_{n-r,k-r} \end{pmatrix}, \tag{19.49}$$

где индексы указывают число строк и столбцов в блоке. Не теряя общности, можно предположить, что $X_{r,r}$ невырождена и, следовательно, имеет обратную матрицу $X_{r,r}^{-1}$. Так как последние $n - r$ строк матрицы X линейно зависят от первых r строк, то $X_{n-r,r} = CX_{r,r}$ и $X_{n-r,k-r} = CX_{r,k-r}$ для некоторой $(n-r) \times r$-матрицы C. Определим теперь новую матрицу порядка $k \times (k-r)$:

$$D = \begin{pmatrix} X_{r,r}^{-1} X_{r,k-r} \\ -I_{k-r} \end{pmatrix}, \tag{19.50}$$

где I_{k-r} — единичная матрица порядка $k - r$. Очевидно, что ранг D равен $k - r$. Вычисляя произведение XD, мы видим, что

$$XD = 0. \tag{19.51}$$

Умножая затем (19.48) справа на $\boldsymbol{D}$, в силу (19.51) получаем

$$\boldsymbol{D} = \boldsymbol{TXD} = \boldsymbol{0}. \tag{19.52}$$

Последнее равенство противоречит тому факту, что ранг $\boldsymbol{D}$ равен $k - r$. Следовательно, (19.48) не может выполняться.

19.14 Чтобы справиться с возникшей трудностью, мы введем систему линейных ограничений

$$\boldsymbol{c} = \boldsymbol{B}\theta, \tag{19.53}$$

где $\boldsymbol{c}$ — вектор констант размерности $(k - r)$ и $\boldsymbol{B}$ — матрица порядка $(k - r) \times k$ ранга $(k - r)$. Мы будем искать оценку вида $\boldsymbol{t} = \boldsymbol{Ly} + \boldsymbol{Nc}$. Условие (19.48) теперь принимает вид

$$\boldsymbol{I} = \boldsymbol{LX} + \boldsymbol{NB}. \tag{19.54}$$

Чтобы избежать противоречия, подобного полученному в (19.52), предположим выполненным следующее условие невырожденности:

$$|\boldsymbol{BD}| \neq 0. \tag{19.55}$$

19.15 Теперь мы можем приступить к нахождению решения по методу НК. Матрица $\boldsymbol{B}$ ранга $(k - r)$ устраняет неполноту ранга матрицы $\boldsymbol{X}$. Фактически мы оперируем с $\boldsymbol{c}$, как с вектором вспомогательных случайных величин, и решаем совместно (19.8) и (19.53) для модифицированной модели

$$\begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{c} \end{pmatrix} = \begin{pmatrix} \boldsymbol{X} \\ \boldsymbol{B} \end{pmatrix} \theta + \begin{pmatrix} \varepsilon \\ \boldsymbol{0} \end{pmatrix}. \tag{19.56}$$

Матрица $\begin{pmatrix} \boldsymbol{X} \\ \boldsymbol{B} \end{pmatrix}' \begin{pmatrix} \boldsymbol{X} \\ \boldsymbol{B} \end{pmatrix} = \boldsymbol{XX}' + \boldsymbol{B}'\boldsymbol{B}$ положительно определена, так как ненулевой вектор $\boldsymbol{d}$ обращает $\boldsymbol{d}'\boldsymbol{X}'\boldsymbol{Xd} = (\boldsymbol{Xd})'\boldsymbol{Xd}$ в нуль только тогда, когда $\boldsymbol{Xd} = \boldsymbol{0}$, откуда следует, что $\boldsymbol{d}$ является столбцом $\boldsymbol{D}$. Но тогда (19.55) дает $\boldsymbol{Bd} \neq \boldsymbol{0}$, так что $\boldsymbol{d}'\boldsymbol{B}'\boldsymbol{Bd} > 0$. Таким образом, $\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{B}'\boldsymbol{B}$ строго положительно определена и поэтому может быть обращена.

(19.56) приводит к решению

$$\hat{\theta} = (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{B}'\boldsymbol{B})^{-1}(\boldsymbol{X}'\boldsymbol{y} + \boldsymbol{B}'\boldsymbol{c}), \tag{19.57}$$

аналогичному (19.12) и являющемуся, как и прежде, несмещенной линейной МД-оценкой для θ. Матрица рассеяния этой оценки ввиду того, что $\boldsymbol{c}$ не является случайной величиной, имеет вид

$$\boldsymbol{V}(\hat{\theta}) = \sigma^2 (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{B}'\boldsymbol{B})^{-1} \boldsymbol{X}'\boldsymbol{X} (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{B}'\boldsymbol{B})^{-1}. \tag{19.58}$$

Матрица $\boldsymbol{B}$ в (19.53) есть произвольная матрица, удовлетворяющая (19.55). В самом деле, если мы заменим $\boldsymbol{B}$ на $\boldsymbol{UB}$, где $\boldsymbol{U}$ —

любая невырожденная матрица порядка $(k-r)\times(k-r)$, то значения (19.57) и (19.58) не изменятся. Поэтому в каждом частном случае можно выбирать $\boldsymbol{B}$ из условий удобства вычислений.

19.16 Упражнение 19.8 показывает, что σ^2 несмещенно оценивается суммой квадратов остатков, деленной на $(n-r)$, если $n > r$.

Чипмэн (1964) подробно рассматривает теорию НК в случае вырождения. См. также Т. Льюис и Оделл (1966).

Пример 19.9

В качестве простого примера вырождения предположим, что

$$\boldsymbol{\theta}=\begin{pmatrix}\theta_1\\ \theta_2\\ \theta_3\end{pmatrix}, \qquad \boldsymbol{X}=\begin{bmatrix}1 & 1 & 0\\ 1 & 0 & 1\\ 1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}.$$

Здесь $n = 4$ и $k = 3$. Матрица $\boldsymbol{X}$ имеет ранг $2 < k$ в силу линейной зависимости между ее столбцами:

$$\boldsymbol{x}_1 - \boldsymbol{x}_2 - \boldsymbol{x}_3 = \boldsymbol{0}.$$

Прежде всего мы видим, что, как было установлено в **19.13**, вектор $\boldsymbol{\theta}$ не может быть оценен несмещенно.

Матрица (19.50) порядка 3×1, выражающая линейную зависимость, имеет вид

$$\boldsymbol{D}=\begin{pmatrix}\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}^{-1}\cdot\begin{pmatrix}0\\ 1\end{pmatrix}\\ -1\end{pmatrix}=\begin{pmatrix}1\\ -1\\ -1\end{pmatrix}.$$

Введем теперь матрицу порядка 1×3

$$\boldsymbol{B}=(1 \quad 0 \quad 0),$$

которая удовлетворяет (19.55), поскольку $\boldsymbol{BD} = 1$ (в рассматриваемом случае одного линейного соотношения это скаляр).

Следовательно, (19.53) имеет вид

$$c=(1 \quad 0 \quad 0)\begin{pmatrix}\theta_1\\ \theta_2\\ \theta_3\end{pmatrix}=\theta_1$$

(снова скаляр). Согласно (19.57) НК-оценка равна

$$\begin{pmatrix}\hat\theta_1\\ \hat\theta_2\\ \hat\theta_3\end{pmatrix}=\left[\begin{pmatrix}1&1&1&1\\1&0&1&0\\0&1&0&1\end{pmatrix}\begin{pmatrix}1&1&0\\1&0&1\\1&1&0\\1&0&1\end{pmatrix}+\begin{pmatrix}1\\0\\0\end{pmatrix}(1\ \ 0\ \ 0)\right]^{-1}\times$$

$$\times\left[\begin{pmatrix}1&1&1&1\\1&0&1&0\\0&1&0&1\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\\y_4\end{pmatrix}+\begin{pmatrix}1\\0\\0\end{pmatrix}c\right]=\begin{pmatrix}5&2&2\\2&2&0\\2&0&2\end{pmatrix}^{-1}\begin{pmatrix}y_1+y_2+y_3+y_4+c\\y_1+y_3\\y_2+y_4\end{pmatrix}=$$

$$=\begin{pmatrix}1&-1&-1\\-1&3/2&1\\-1&1&3/2\end{pmatrix}\begin{pmatrix}\sum y+c\\y_1+y_3\\y_2+y_4\end{pmatrix}=\begin{pmatrix}c\\(1/2)(y_1+y_3)-c\\(1/2)(y_2+y_4)-c\end{pmatrix}.$$

Поскольку мы выбрали $\boldsymbol{B}$ так, что $c=\theta_1$, то мы не можем получить ничего лучшего для оценки самого θ. Однако в силу (19.9) любая система линейных функций $\boldsymbol{C}\theta$ несмещенно оценивается статистикой $\boldsymbol{Ty}$, если $\boldsymbol{TX}=\boldsymbol{C}$. В данном случае несмещенно оцениваются $(\theta_1+\theta_2)$ и $(\theta_1+\theta_3)$, так как $\boldsymbol{C}=\begin{pmatrix}1&1&0\\1&0&1\end{pmatrix}$ удовлетворяет (19.19) при $\boldsymbol{T}=\begin{pmatrix}1/2&0&1/2&0\\0&1/2&0&1/2\end{pmatrix}$. Оценками для $(\theta_1+\theta_2)$ и $(\theta_1+\theta_3)$ служат, следовательно, $(1/2)(y_1+y_3)$ и $(1/2)(y_2+y_4)$.

Из (19.58) получаем

$$\boldsymbol{V}(\hat\theta)=\sigma^2\begin{pmatrix}1&-1&-1\\-1&3/2&1\\-1&1&3/2\end{pmatrix}\begin{pmatrix}4&2&2\\2&2&0\\2&0&2\end{pmatrix}\begin{pmatrix}1&-1&-1\\-1&3/2&1\\-1&1&3/2\end{pmatrix}=$$

$$=\sigma^2\begin{pmatrix}0&0&0\\0&1/2&0\\0&0&1/2\end{pmatrix},$$

так что

$$\boldsymbol{D}(\hat\theta_1+\hat\theta_2)=\boldsymbol{D}\hat\theta_2=\boldsymbol{D}(\hat\theta_1+\hat\theta_3)=\boldsymbol{D}\hat\theta_3=\sigma^2/2,$$

что, впрочем, непосредственно следует из того факта, что каждая из оценок равна среднему двух наблюдений с дисперсией σ^2. Кроме того,

$$\operatorname{cov}(\hat\theta_1+\hat\theta_2,\ \hat\theta_1+\hat\theta_3)=0.$$

Это полезное свойство следует из ортогональности второго и третьего столбцов X. При обсуждении приложений теории НК к дисперсионному анализу в 3-м томе мы снова вернемся к этой теме.

Более общая линейная модель

19.17 В развитой нами теории НК всюду предполагалось, что выполнено (19.10), т. е. что ошибки некоррелированы и имеют постоянную дисперсию. Не представляет труда обобщить эту линейную модель на более общий случай, когда матрица рассеяния ошибок равна $\sigma^2 \boldsymbol{V}$, где $\boldsymbol{V}$ невырождена. В этом случае (подробное рассмотрение предоставляется читателю — см. упражнения 19.2 и 19.5) мы получаем, что (19.24) переходит в

$$\boldsymbol{t} = \boldsymbol{C}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} \tag{19.59}$$

и что (19.59) дает линейную несмещенную МД-оценку для $\boldsymbol{C\theta}$. Кроме того, (19.26) принимает вид

$$\boldsymbol{V}(\boldsymbol{t}) = \sigma^2 \boldsymbol{C}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{C}'. \tag{19.60}$$

В частности, если $\boldsymbol{V}$ диагональна, но не равна $\boldsymbol{I}$, т. е. ε_i некоррелированы, но имеют различные дисперсии, то требуемая система оценок дается формулой (19.59).

Конечно, чтобы использовать эти уравнения, мы должны знать $\boldsymbol{V}$. На практике, однако, $\boldsymbol{V}$ часто неизвестно, что делает задачу оценивания гораздо более трудной.

Порядковые оценки наименьших квадратов для параметров расположения и масштаба

19.18 Одной из ситуаций, в которой (19.59) и (19.60) оказываются полезными, является оценивание параметров расположения и масштаба, исходя из *порядковых статистик*, т. е. выборочных наблюдений, упорядоченных по величине. Излагаемые ниже результаты принадлежат Ллойду (1952) и Даунтону (1953).

Как и раньше (см. **14.1**), $y_{(1)}, y_{(2)}, \ldots, y_{(n)}$ обозначают порядковые статистики, μ и σ, как обычно, — параметры расположения и масштаба, подлежащие оцениванию (μ и σ не обязательно являются средним значением и стандартным отклонением распределения). Введем также обозначение

$$z_{(r)} = (y_{(r)} - \mu)/\sigma, \qquad r = 1, 2, \ldots, n. \tag{19.61}$$

Пусть

$$\mathbf{M}(\boldsymbol{z}) = \boldsymbol{\alpha}, \qquad \boldsymbol{V}(\boldsymbol{z}) = \boldsymbol{V}, \tag{19.62}$$

где $\boldsymbol{z}$ — вектор-столбец размерности n с компонентами $z_{(r)}$. Поскольку $\boldsymbol{z}$ нормировано соотношением (19.61), то $\boldsymbol{\alpha}$ и $\boldsymbol{V}$ не зависят от μ и σ.

Далее, из (19.61) и (19.62) получаем

$$\mathbf{M}(\boldsymbol{y}) = \mu\mathbf{1} + \sigma\boldsymbol{\alpha}, \tag{19.63}$$

где $\boldsymbol{y}$ — вектор с компонентами $y_{(r)}$ и $\mathbf{1}$ — вектор из единиц, и кроме того,

$$\boldsymbol{V}(\boldsymbol{y}) = \sigma^2 \boldsymbol{V}. \tag{19.64}$$

Мы можем теперь, применяя (19.59) и (19.60), найти НК-оценки для μ и σ. Получаем

$$\begin{pmatrix} \hat{\mu} \\ \hat{\sigma} \end{pmatrix} = \{(\mathbf{1} \vdots \boldsymbol{\alpha})' \boldsymbol{V}^{-1} (\mathbf{1} \vdots \boldsymbol{\alpha})\}^{-1} (\mathbf{1} \vdots \boldsymbol{\alpha})' \boldsymbol{V}^{-1} \boldsymbol{y} \tag{19.65}$$

и

$$\boldsymbol{V}\begin{pmatrix} \hat{\mu} \\ \hat{\sigma} \end{pmatrix} = \sigma^2 \{(\mathbf{1} \vdots \boldsymbol{\alpha})' \boldsymbol{V}^{-1} (\mathbf{1} \vdots \boldsymbol{\alpha})\}^{-1}. \tag{19.66}$$

Далее,

$$\{(\mathbf{1} \vdots \boldsymbol{\alpha})' \boldsymbol{V}^{-1} (\mathbf{1} \vdots \boldsymbol{\alpha})\}^{-1} = \begin{pmatrix} \mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1} & \mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{\alpha} \\ \mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{\alpha} & \boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{\alpha} \end{pmatrix}^{-1} =$$

$$= \frac{1}{\Delta}\begin{pmatrix} \boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{\alpha} & -\mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{\alpha} \\ -\mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{\alpha} & \mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1} \end{pmatrix}, \tag{19.67}$$

где

$$\Delta = \{(\mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1})(\boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{\alpha}) - (\mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{\alpha})^2\}. \tag{19.68}$$

Из (19.65) и (19.67) находим выражения для оценок μ и σ:

$$\left.\begin{aligned} \hat{\mu} &= -\boldsymbol{\alpha}'\boldsymbol{V}^{-1}(\mathbf{1}\boldsymbol{\alpha}' - \boldsymbol{\alpha}\mathbf{1}')\boldsymbol{V}^{-1}\boldsymbol{y}/\Delta, \\ \hat{\sigma} &= \mathbf{1}'\boldsymbol{V}^{-1}(\mathbf{1}\boldsymbol{\alpha}' - \boldsymbol{\alpha}\mathbf{1}')\boldsymbol{V}^{-1}\boldsymbol{y}/\Delta, \end{aligned}\right\} \tag{19.69}$$

а из (19.66) и (19.67) — дисперсии и ковариацию этих оценок:

$$\left.\begin{aligned} \mathrm{D}\hat{\mu} = \sigma^2\boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{\alpha}/\Delta, \quad \mathrm{D}\hat{\sigma} = \sigma^2\mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1}/\Delta, \\ \mathrm{cov}\,(\hat{\mu}, \hat{\sigma}) = -\sigma^2\mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{\alpha}/\Delta. \end{aligned}\right\} \tag{19.70}$$

19.19 Поскольку $\boldsymbol{V}$ и $\boldsymbol{V}^{-1}$ положительно определены, то можно написать

$$\left.\begin{aligned} \boldsymbol{V} &= \boldsymbol{T}\boldsymbol{T}', \\ \boldsymbol{V}^{-1} &= (\boldsymbol{T}^{-1})'\boldsymbol{T}^{-1}, \end{aligned}\right\} \tag{19.71}$$

так что для произвольного вектора $\boldsymbol{b}$ имеем

$$\boldsymbol{b}'\boldsymbol{V}\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{T}\boldsymbol{T}'\boldsymbol{b} = (\boldsymbol{T}'\boldsymbol{b})'(\boldsymbol{T}'\boldsymbol{b}) = \sum_{i=1}^{n} h_i^2,$$

где h_i есть i-й элемент вектора $\boldsymbol{T}'\boldsymbol{b}$.

Аналогично для вектора $\boldsymbol{c}$ получаем

$$\boldsymbol{c}'\boldsymbol{V}^{-1}\boldsymbol{c} = (\boldsymbol{T}^{-1}\boldsymbol{c})'(\boldsymbol{T}^{-1}\boldsymbol{c}) = \sum_{i=1}^{n} k_i^2,$$

где k_i есть i-й элемент $\boldsymbol{T}^{-1}\boldsymbol{c}$. Пользуясь неравенством Коши, находим

$$\sum h_i^2 \sum k_i^2 = \boldsymbol{b}'\boldsymbol{V}\boldsymbol{b} \cdot \boldsymbol{c}'\boldsymbol{V}^{-1}\boldsymbol{c} \geqslant \left(\sum h_i k_i\right)^2 = \left\{(\boldsymbol{T}'\boldsymbol{b})'(\boldsymbol{T}^{-1}\boldsymbol{c})\right\}^2 = (\boldsymbol{b}'\boldsymbol{c})^2. \tag{19.72}$$

Полагая в (19.72)

$$\boldsymbol{b} = (\boldsymbol{V}^{-1} - \boldsymbol{I})\,\mathbf{1}, \qquad \boldsymbol{c} = \boldsymbol{\alpha}, \tag{19.73}$$

получаем неравенство

$$\mathbf{1}'(\boldsymbol{V}^{-1} - \boldsymbol{I})\boldsymbol{V}(\boldsymbol{V}^{-1} - \boldsymbol{I})\,\mathbf{1} \cdot \boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{\alpha} \geqslant \left\{\mathbf{1}'(\boldsymbol{V}^{-1} - \boldsymbol{I})\,\boldsymbol{\alpha}\right\}^2. \tag{19.74}$$

Поскольку, как нетрудно проверить *),

$$n = \mathbf{1}'\mathbf{1} = \mathbf{1}'\boldsymbol{V}\mathbf{1}, \qquad \mathbf{1}'\boldsymbol{\alpha} = 0, \tag{19.75}$$

то (19.74) переписывается в виде

$$(\mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1} - n)\,\boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{\alpha} \geqslant (\mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{\alpha})^2,$$

откуда, используя (19.70) и (19.68) и считая теперь σ^2 дисперсией, получаем

$$\mathrm{D}\hat{\mu} \leqslant \sigma^2/n = \mathrm{D}\bar{y}. \tag{19.76}$$

Соотношение (19.76) довольно очевидно: выборочное среднее $\bar{y}$, будучи линейной оценкой, не может иметь дисперсию меньшую, чем МД-оценка $\hat{\mu}$. Приведенное выше доказательство, однако, позволяет нам определить условия, при которых (19.76) превращается в равенство. Это происходит, когда неравенство Коши (19.72) становится равенством, т. е. когда $h_i = \lambda k_i$ для некоторого постоянного λ, или

$$\boldsymbol{T}'\boldsymbol{b} = \lambda \boldsymbol{T}^{-1}\boldsymbol{c}.$$

Вследствие (19.73) это условие записывается в виде

$$\boldsymbol{T}'(\boldsymbol{V}^{-1} - \boldsymbol{I})\,\mathbf{1} = \lambda \boldsymbol{T}^{-1}\boldsymbol{\alpha}$$

или

$$\boldsymbol{T}\boldsymbol{T}'(\boldsymbol{V}^{-1} - \boldsymbol{I})\,\mathbf{1} = \lambda\boldsymbol{\alpha}. \tag{19.77}$$

Отсюда, используя (19.71), получаем окончательно следующее условие выполнения равенства $\mathrm{D}\hat{\mu} = \mathrm{D}\bar{y} = \sigma^2/n$:

$$(\boldsymbol{I} - \boldsymbol{V})\,\mathbf{1} = \lambda\boldsymbol{\alpha}. \tag{19.78}$$

Если (19.78) выполнено, то в силу единственности решения уравнений метода НК получим

$$\hat{\mu} = \bar{y}, \tag{19.79}$$

что можно проверить, применяя (19.78) к $\hat{\mu}$ в (19.69).

*) Равенство (19.75) получается в предположении, что μ и σ^2 — среднее и дисперсия исходного распределения. (*Прим. ред.*)

19.20 В случае симметричного исходного распределения ситуация упрощается, так как компоненты вектора математических ожиданий

$$\boldsymbol{\alpha} = \begin{bmatrix} \mathsf{M}(z_{(1)}) \\ \cdot \\ \cdot \\ \cdot \\ \mathsf{M}(z_{(n)}) \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_n \end{bmatrix}$$

удовлетворяют условию *)

$$\alpha_i = -\alpha_{n+1-i} \quad \text{(для всех } i\text{)}, \tag{19.80}$$

непосредственно вытекающему из (14.2). Следовательно,

$$\boldsymbol{\alpha}'\boldsymbol{V}^{-1}\mathbf{1} = \mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{\alpha} = 0 \tag{19.81}$$

и (19.69) принимает вид

$$\left.\begin{aligned} \hat{\mu} &= \mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{y}/\mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1}, \\ \hat{\sigma} &= \boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{y}/\boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{\alpha}, \end{aligned}\right\} \tag{19.82}$$

а (19.70) упрощается следующим образом:

$$\left.\begin{aligned} \mathsf{D}\hat{\mu} = \sigma^2/\mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1}, \quad \mathsf{D}\hat{\sigma} = \sigma^2/\boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{\alpha}, \\ \operatorname{cov}(\hat{\mu}, \hat{\sigma}) = 0. \end{aligned}\right\} \tag{19.83}$$

Таким образом, порядковые НК-оценки $\hat{\mu}$ и $\hat{\sigma}$ некоррелированы, если исходное распределение симметрично, — результат, аналогичный полученному в **18.34** для МП-оценок.

Пример 19.10

Оценить средину размаха (или среднее) μ и размах σ прямоугольного распределения

$$dF(x) = \frac{dx}{\sigma}, \qquad \mu - \tfrac{1}{2}\sigma \leqslant x \leqslant \mu + \tfrac{1}{2}\sigma.$$

Производя нормировку (19.61) и учитывая (14.2), легко показать, что

$$\alpha_r = \mathsf{M}(z_{(r)}) = \frac{r}{n+1} - \frac{1}{2} \tag{19.84}$$

и что элементы матрицы рассеяния $\boldsymbol{V}$ величин $z_{(r)}$ равны

$$V_{rs} = r(n-s+1)/\{(n+1)^2(n-2)\}, \qquad r \leqslant s. \tag{19.85}$$

*) Предполагается, что μ — среднее значение исходного распределения. (*Прим. ред.*)

Обратная матрица для $\boldsymbol{V}$ имеет вид

$$\boldsymbol{V}^{-1} = (n+1)(n+2)\begin{Vmatrix} 2 & -1 & & & & \\ & \cdot & \cdot & & 0 & \\ -1 & & 2 & \cdot & & \\ & \cdot & & \cdot & \cdot & -1 \\ & & \cdot & & \cdot & \\ & 0 & & \cdot & & \cdot \\ & & & & -1 & 2 \end{Vmatrix}, \tag{19.86}$$

откуда получаем

$$\mathbf{1}'\boldsymbol{V}^{-1} = (n+1)(n+2)\begin{Vmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \\ 0 \\ 1 \end{Vmatrix} \tag{19.87}$$

и, учитывая (19.84),

$$\boldsymbol{\alpha}'\boldsymbol{V}^{-1} = \tfrac{1}{2}(n+1)(n+2)\begin{Vmatrix} -1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \\ 0 \\ 1 \end{Vmatrix}. \tag{19.88}$$

В силу (19.87) и (19.88) формулы (19.82) и (19.83) дают

$$\left.\begin{gathered} \hat{\mu} = \tfrac{1}{2}(y_{(1)} + y_{(n)}), \quad \hat{\sigma} = (n+1)(y_{(n)} - y_{(1)})/(n-1), \\ \mathrm{D}\hat{\mu} = \sigma^2/\{2(n+1)(n+2)\}, \quad \mathrm{D}\hat{\sigma} = 2\sigma^2/\{(n-1)(n+2)\}, \\ \operatorname{cov}(\hat{\mu}, \hat{\sigma}) = 0. \end{gathered}\right\} \tag{19.89}$$

За исключением поправки на смещение для $\hat{\sigma}$, результаты совпадают с теми, которые были получены в примере 18.12 методом МП. Этого, однако, следовало ожидать, так как $y_{(1)}$ и $y_{(n)}$ образуют пару совместно достаточных статистик для μ и σ (пример 17.21).

19.21 Из примера 19.10 видно, что для использования теории, развитой в **19.18—20**, мы должны определить матрицу рассеяния $\boldsymbol{V}$ нормированных порядковых статистик, зависящую от вида исходного распределения. Это является полной противоположностью теории НК, рассмотренной ранее в этой главе, в

которой не предполагается знания вида исходного распределения. В главе 32 мы возвратимся к рассмотрению свойств порядковых статистик, используемых для оценивания параметров.

Общая теория НК, развитая в этой главе, играет основную роль во многих разделах статистической теории, и мы будем неоднократно использовать полученные результаты в последующих главах.

Другие методы оценивания

19.22 В предыдущей главе мы видели, что, кроме того факта, что МП-оценки являются функциями от достаточных статистик для оцениваемых параметров, все остальные свойства этих оценок носят асимптотический характер, именно:

(I) состоятельность;

(II) асимптотическая нормальность;

(III) несмещенность.

Очевидно, МП-оценка $\hat{\theta}$ не может быть единственной оценкой, обладающей этими свойствами. Например, добавление к $\hat{\theta}$ произвольной константы c/n^r не изменяет ее свойств первого порядка, если r достаточно велико. Поэтому естественно вслед за Нейманом (1949) исследовать класс оценок, обладающих теми же асимптотическими свойствами, что и $\hat{\theta}$. Дополнительным стимулом к такому исследованию является еще то, что иногда нахождение МП-оценок связано с большими вычислительными трудностями (см. примеры 18.3, 18.9, 18.10).

19.23 Предположим, что имеется s ($\geqslant 1$) выборок с n_i наблюдениями в i-й выборке. Как и в **18.19**, упростим задачу, предполагая, что каждое наблюдение в i-й выборке относится к одному из k_i взаимно исключающих классов. Если π_{ij} обозначает вероятность попадания в j-й класс наблюдения из i-й выборки, то

$$\sum_{j=1}^{k_i} \pi_{ij} = 1, \tag{19.90}$$

и задача сводится к рассмотрению системы из s мультиномиальных распределений. Пусть n_{ij} — число наблюдений из j-й выборки, попавших в i-й класс, и $p_{ij} = n_{ij}/n_i$ — соответствующая относительная частота. Вероятности π_{ij} являются функциями от неизвестных параметров $(\theta_1, \ldots, \theta_r)$.

Функция T от случайных величин называется наилучшей асимптотически нормальной оценкой (сокращенно: НАН-оценкой) для одного из неизвестных параметров θ_1, если

(I) $T(\{p_{ij}\})$ состоятельна для θ_1;

(II) T асимптотически нормальна при $N = \sum_{i=1}^{s} n_i \to \infty$;

(III) T эффективна;

(IV) $\partial T/\partial p_{ij}$ существует и непрерывна по p_{ij} для всех i, j.

Первые три условия в точности совпадают с теми, которые мы уже доказали для МП-оценок в главе 18. Легко проверить, что в рассматриваемой мультиномиальной ситуации для МП-оценок выполнено также четвертое свойство. Таким образом, класс НАН-оценок содержит МП-оценку как частный случай.

19.24 Нейманом было показано, что следующие три условия являются необходимыми и достаточными для того, чтобы некоторая оценка была НАН:

(I) $T(\{\pi_{ij}\}) \equiv \theta_1$;

(II) выполнено условие (IV) пункта **19.23**;

(III) $\sum_{i=1}^{s} \frac{1}{n_i} \sum_{j=1}^{k_i} \left[\left(\frac{\partial T}{\partial p_{ij}} \right)_{p_{ij}=\pi_{ij}} \right]^2 \pi_{ij}$ обращается в минимум по $\partial T/\partial p_{ij}$.

Условие (I) достаточно для состоятельности, и, вообще говоря, оно является более сильным условием, чем состоятельность *). В нашем случае, поскольку статистика T есть непрерывная функция от p_{ij}, а p_{ij} сходятся по вероятности к π_{ij}, T сходится по вероятности к $T(\{\pi_{ij}\})$, т. е. к θ_1.

Условие (III) есть просто условие эффективности, так как подлежащая минимизации функция является дисперсией T, к которой применено необходимое условие минимума

$$\sum_{j} \left(\frac{\partial T}{\partial p_{ij}} \right)_{p_{ij}=\pi_{ij}} \cdot \pi_{ij} = 0.$$

В том виде, как они здесь сформулированы, эти три условия не имеют большой ценности. Однако Нейман показал, что можно получить достаточную систему условий, если заменить (III) условием непосредственно на $\partial T/\partial p_{ij}$, которое мы здесь не приводим. Из этих условий он выводит, что

(а) МП-оценка является НАН-оценкой, что мы уже видели;

(б) оценки, полученные по так называемому методу минимума хи-квадрат, также являются НАН-оценками.

Теперь мы приступаем к изучению этого второго класса оценок.

Оценивание по методу минимума хи-квадрат

19.25 Возвращаясь к ситуации, рассмотренной в **19.23**, назовем статистику T оценкой по методу минимума хи-квадрат

*) В действительности первоначально определение состоятельности было дано Фишером (1921а) именно в форме (I).

(МХК-оценкой), если она получена минимизацией по θ_1 выражения

$$\chi^2 = \sum_{i=1}^{s} \frac{1}{n_i} \sum_{j=1}^{k_i} \frac{(p_{ij} - \pi_{ij})^2}{\pi_{ij}} = \sum_{i=1}^{s} \frac{1}{n_i} \left(\sum_{j=1}^{k_i} \frac{p_{ij}^2}{\pi_{ij}} - 1 \right), \quad (19.91)$$

где π_{ij} — функции от $\theta_1, \ldots, \theta_r$. Минимизация (19.91) приводит к уравнению

$$\frac{\partial \chi^2}{\partial \theta_1} = -\sum_i \frac{1}{n_i} \sum_j \left(\frac{p_{ij}}{\pi_{ij}} \right)^2 \frac{\partial \pi_{ij}}{\partial \theta_1} = 0. \quad (19.92)$$

Корень (19.92), рассматриваемого как уравнение относительно θ_1, будет МХК-оценкой для θ_1. Очевидно, можно, обобщая (19.92), получить систему из r уравнений, которые надо решить совместно, чтобы найти МХК-оценки для $\theta_1, \ldots, \theta_r$.

Процедура нахождения МХК-оценок совершенно аналогична процедуре нахождения МП-оценок, рассмотренной в главе 18. Более того, (асимптотические) свойства МХК- и МП-оценок очень похожи. Действительно, с вероятностью 1 имеется единственный состоятельный корень уравнений МХК и этот корень соответствует абсолютному минимуму (нижней грани) (19.91). Для простейшего случая $s = 1$ доказательство этого факта дано С. Р. Рао (1957).

19.26 Видоизмененная форма МХК-оценок получается, если минимизировать

$$(\chi')^2 = \sum_{i=1}^{s} \frac{1}{n_i} \sum_{j=1}^{k_i} \frac{(p_{ij} - \pi_{ij})^2}{p_{ij}} = \sum_i \frac{1}{n_i} \left(\sum_j \frac{\pi_{ij}^2}{p_{ij}} - 1 \right) \quad (19.93)$$

вместо (19.91). В (19.93) предполагается, что p_{ij} не обращаются в нуль. Чтобы минимизировать (19.93), надо решить относительно θ_1 уравнение

$$\frac{\partial (\chi')^2}{\partial \theta_1} = 2 \sum_i \frac{1}{n_i} \sum_j \left(\frac{\pi_{ij}}{p_{ij}} \right) \frac{\partial \pi_{ij}}{\partial \theta_1} = 0. \quad (19.94)$$

Такая модификация МХК-оценок, как показал Нейман (1949), также является НАН-оценкой.

19.27 Поскольку методы МП, МХК и видоизмененный метод МХК имеют одинаковые асимптотические свойства, то при выборе между ними в любом конкретном случае надо исходить либо из удобства вычислений, либо из выборочных свойств при малых выборках, либо из того и другого. Что касается вычислительных удобств, то в общем случае об этом можно сказать немногое. Иногда труднее решать уравнения МП, а иногда —

уравнения МХК. Но если мы имеем дело с непрерывным распределением, то для применения метода МХК необходимо сгруппировать наблюдения. Кажется, однако, нежелательным производить группировку, нужную лишь для оценивания и бесполезную в других отношениях. Кроме того, иногда трудно, особенно в случае непрерывных распределений, выразить π_{ij} в терминах оцениваемых параметров. Наше мнение, следовательно, состоит в том, что имеющая в настоящее время место тенденция к оцениванию по методу МП в достаточно широком классе случаев оправдана с точки зрения вычислений. Следующий пример иллюстрирует вычислительную процедуру метода МХК в относительно простом случае.

Пример 19.11

Рассмотрим оценивание параметра θ распределения Пуассона по одной выборке из n наблюдений. Мы знаем (примеры 17.8, 17.15), что выборочное среднее $\bar{x}$ есть достаточная МГД-оценка для θ, а из **18.5** следует, что $\bar{x}$ есть также МП-оценка. МХК-оценка для θ, однако, не равна $\bar{x}$, что иллюстрирует тот факт, что метод МХК не обязательно приводит к одномерной достаточной статистике, если последняя существует.

Теоретические вероятности в данном случае равны

$$\pi_j = e^{-\theta}\theta^j/j!, \qquad j = 0,\ 1,\ 2,\ \ldots,$$

откуда

$$\frac{\partial \pi_j}{\partial \theta} = \pi_j\left(\frac{j}{\theta} - 1\right).$$

Уравнения минимизации (19.92) можно, следовательно, опуская множитель $1/n$, записать в виде

$$-\sum\left(\frac{p_i}{\pi_j}\right)^2 \pi_j\left(\frac{j}{\theta} - 1\right) = \sum_j \frac{p_j^2}{\pi_j}\left(1 - \frac{j}{\theta}\right) = 0. \qquad (19.95)$$

Для решения этого уравнения относительно θ мы используем итерационный метод, подобный тому, который был использован для нахождения МП-оценки в **18.21**. Разложим левую часть (19.95) в ряд Тейлора, как функцию от θ в точке $\bar{x}$ (выборочное среднее), рассматриваемой как начальное значение. С точностью до первого порядка получаем

$$\sum_j \frac{p_j^2}{\pi_j}\left(1 - \frac{j}{\theta}\right) = \sum_j \frac{p_j^2}{m_j}\left(1 - \frac{j}{\bar{x}}\right) + (\theta - \bar{x})\sum_j \frac{p_j^2}{m_j}\left\{\frac{j}{\bar{x}^2} + \left(1 - \frac{j}{\bar{x}}\right)^2\right\}, \qquad (19.96)$$

где использовано обозначение $m_j = e^{-\bar{x}}\bar{x}^j/j!$ Приравнивая (19.96) к нулю и учитывая (19.95), находим

$$(\theta - \bar{x}) = \bar{x} \cdot \frac{\sum_j \frac{p_j^2}{m_j}(j - \bar{x})}{\sum_j \frac{p_j^2}{m_j}\{j + (j - \bar{x})^2\}}. \qquad (19.97)$$

Используя (19.97), мы получаем улучшенное значение θ и затем повторяем этот процесс столько раз, сколько нужно.

В качестве численного примера мы используем данные Уайтекера (1914) о числе смертей женщин в возрасте старше 85 лет. Эти данные составлены из ежедневных сообщений газеты «Таймс» в 1910—1912 гг. и охватывают период в 1096 дней. Соответствующее распределение приведено в первых двух столбцах таблицы. Среднее число смертей равно $\bar{x} = 1295/1096 = 1{,}181569$. Это и есть значение МП-оценки, которое мы возьмем в качестве начального приближения для МХК-оценки. Третий столбец дает ожидаемые частоты пуассоновского распределения с параметром $\bar{x}$. В оставшихся пяти столбцах приведены вычисления, относящиеся к (19.97).

Число смертей (j)	Частота (np_j)	nm_j	$\frac{np_j}{nm_j} - \frac{np_j^2}{m_j}$	$(j-\bar{x})$	$\frac{np_j^2}{m_j}(j-\bar{x})$	$\{j+(j-\bar{x})^2\}$	$\frac{np_j^2}{m_j}\{j+(j-\bar{x})^2\}$
0	364	336,25	394,1	−1,1816	−465,7	1,396	551,1
1	376	397,30	355,8	−0,1816	−64,6	1,033	365,9
2	218	234,72	202,5	0,8184	165,8	2,670	540,6
3	89	92,45	85,69	1,8184	155,8	6,307	540,4
4	33	27,31	39,87	2,8184	112,4	11,943	476,1
5	13	6,45	26,20	3,8184	100,0	19,580	512,9
6	2	1,27	3,15	4,8184	15,2	29,217	92,0
7	1	0,25	4,00	5,8184	23,3	40,854	163,4
Всего	n=1096	1096,00			+42,2		3242,4

В результате мы получаем из (19.97) улучшенное значение:

$$\theta = 1{,}1816\left\{1 + \frac{42{,}2}{3242{,}4}\right\} = 1{,}198.$$

К. Смит (1916) получил значение 1,196903, используя большее число итераций.

Смит приводит также детали вычислительной процедуры для оценивания параметров непрерывного распределения, которая в этом случае значительно более трудоемка.

19.28 Вторым критерием для выбора между МП- и МХК-оценками служат свойства, относящиеся к малым выборкам, которые в большей степени пригодны для общего исследования, чем вычислительные аспекты. С. Р. Рао (1961, 1962а) ввел понятие эффективности второго порядка и показал, что в мультиномиальной модели пункта **18.19** при условиях регулярности среди всех НАН-оценок только МП-оценки являются оптимальными в смысле эффективности второго порядка. Берксон (1955, 1956) осуществил выборочные эксперименты, которые показали, что в ситуациях, возникающих в биологических испытаниях, использование МП-оценки иногда приводит к трудности, связанной с тем, что МП-оценка обращается в бесконечность, тогда как некоторая другая НАН-оценка имеет меньшую среднеквадратичную ошибку. При чтении этих статей следует, однако, обратиться также к статье Силверстоуна (1957), который указывает на содержащиеся в них ошибки.

УПРАЖНЕНИЯ

19.1 Предположим, что в линейной модели (19.8) введен еще один параметр θ_0, так что получается новая модель

$$y = X\theta + 1\theta_0 + \varepsilon,$$

где $\mathbf{1}$ — вектор-столбец из единиц размерности n. Показать, что в новой модели НК-оценка для первоначального вектора θ из k параметров имеет в точности ту же форму (19.12), что и в первоначальной модели, с той лишь разницей, что y_i заменены на $(y_i - \bar{y})$ и x_{ij} — на $(x_{ij} - \bar{x}_j)$ для $i = 1, 2, \ldots, n$ и $j = 1, 2, \ldots, k$.

19.2 В случае, когда в линейной модели (19.8) простая матрица рассеяния (19.10) заменена невырожденной матрицей $\sigma^2 V$, допускающей корреляции и неравные дисперсии ε_i, показать, полагая $w = T'y$, где $TT' = V^{-1}$, что НК-оценка, являющаяся несмещенной для $C\theta$, имеет вид

$$C\hat{\theta} = C(X'V^{-1}X)^{-1}X'V^{-1}y.$$

(См. Эйткин, 1935; Плэкетт, 1949.)

19.3 Обобщая (19.38), показать, что если $\mathsf{M}(\varepsilon\varepsilon') = \sigma^2 V$, то $\mathsf{M}(\varepsilon' B\varepsilon) = \sigma^2 \operatorname{tr}(BV)$. Показать далее, что $\mathsf{D}(\varepsilon' B\varepsilon) = 2\sigma^4 \operatorname{tr}(BVBV)$, если ε имеет многомерное нормальное распределение.

19.4 Показать, что в **19.12** отношение $(X\hat{\theta})'(X\hat{\theta})/(ks^2)$ имеет распределение Фишера F с k и $n-k$ степенями свободы, если $\theta = 0$.

19.5 В упражнении 19.2 показать, что обобщение (19.26) имеет вид

$$V(C\hat{\theta}) = \sigma^2 C(X'V^{-1}X)^{-1}C',$$

а обобщением (19.40) в силу упражнения 19.3 является

$$\mathsf{M}\{\varepsilon'[V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}]\varepsilon\} = (n-k)\sigma^2.$$

19.6 Доказать утверждение пункта **19.15** о том, что $\hat{\theta}$ и $V(\hat{\theta})$ в случае вырождения не изменяются при замене B на UB, где U невырождена.

19.7 Используя (19.51), (19.54) и (19.55), показать, что

$$(\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{B}'\boldsymbol{B})^{-1}\,\boldsymbol{B}'\boldsymbol{B} = \boldsymbol{D}\,(\boldsymbol{B}\boldsymbol{D})^{-1}\,\boldsymbol{B}$$

и что, следовательно, (19.58) можно записать в виде

$$\boldsymbol{V}\,(\hat{\theta})\,(\boldsymbol{X}'\boldsymbol{X}) = \sigma^2\{\boldsymbol{I}_k - \boldsymbol{D}\,(\boldsymbol{B}\boldsymbol{D})^{-1}\,\boldsymbol{B}\}.$$

(Плэкетт, 1950.)

19.8 Используя результат упражнения 19.7

$$(\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{B}'\boldsymbol{B})^{-1}\,\boldsymbol{B}'\boldsymbol{B} = \boldsymbol{D}\,(\boldsymbol{B}\boldsymbol{D})^{-1}\,\boldsymbol{B}$$

и модифицируя доказательство пункта **19.9**, показать, что несмещенная оценка для σ^2 в случае вырождения имеет вид

$$\frac{1}{n-r}\,(y - \boldsymbol{X}\hat{\theta})'\,(y - \boldsymbol{X}\hat{\theta}).$$

(Плэкетт, 1950.)

19.9 Показать, используя **19.9**, что в случае линейной модели (19.8)—(19.10) статистика $\hat{\theta}'\boldsymbol{A}\hat{\theta} - s^2\,\mathrm{tr}\,\{(\boldsymbol{X}'\boldsymbol{X})^{-1}\,\boldsymbol{A}\}$ является несмещенной оценкой для квадратичной формы $\theta'\boldsymbol{A}\theta$.

19.10 Показать, что в случае симметричного исходного распределения условие того, что порядковая НК-оценка $\hat{\mu}$ в (19.82) равна выборочному среднему $\bar{y} = \mathbf{1}'y/\mathbf{1}'\mathbf{1}$, имеет вид

$$\boldsymbol{V}\mathbf{1} = \mathbf{1},$$

т. е. сумма элементов каждой строки матрицы рассеяния должна быть равна единице. Показать, что это свойство выполнено для одномерного нормального распределения.

(Ллойд, 1952.)

19.11 Для экспоненциального распределения

$$dF\,(y) = \exp\left\{-\left(\frac{y-\mu}{\sigma}\right)\right\}dy/\sigma, \quad \sigma > 0, \quad \mu \leqslant y \leqslant \infty,$$

показать, что элементы $\boldsymbol{\alpha}$ в (19.62) имеют вид

$$\alpha_r = \sum_{i=1}^{r} (n - i + 1)^{-1},$$

а элементы $\boldsymbol{V}$ равны

$$V_{rs} = \sum_{i=1}^{m} (n - i + 1)^{-2}, \quad \text{где} \quad m = \min\,(r, s).$$

На основе этого проверить, что обратная матрица равна

$$\boldsymbol{V}^{-1} = \begin{Bmatrix} n^2 + (n-1)^2 & -(n-1)^2 & & & & \\ -(n-1)^2 & (n-1)^2 + (n-2)^2 & -(n-2)^2 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & & 2^2 & 2^2 + 1^2 & -1^2 \\ & & & & & -1^2 & 1^2 \end{Bmatrix}.$$

19.12 В упражнении 19.11 показать, используя (19.69), что несмещенные МД-оценки даются формулами

$$\hat{\mu} = y_{(1)} - (\bar{y} - y_{(1)})/(n-1), \qquad \hat{\sigma} = n(\bar{y} - y_{(1)})/(n-1).$$

Сравнить эти оценки с МП-оценками тех же параметров.

(Сархан, 1954.)

19.13 Показать, что, когда все n_i велики, минимизация выражения (19.91) или (19.93) для χ^2 дает ту же оценку, что и метод МП.

19.14 Для случая $s = 1$ показать, что первые два момента статистики (19.91) даются выражениями

$$n_1^2 \mathbf{M}(\chi^2) = k_1 - 1,$$

$$n_1^4 \mathbf{D}(\chi^2) = 2(k_1 - 1)\left(1 - \frac{1}{n_1}\right) + \frac{1}{n_1}\sum_{j=1}^{k_1}\frac{1}{\pi_{ij}} - \frac{k_1^2}{n_1}$$

и что для любого $c > 0$ обобщение (19.93) имеет математическое ожидание

$$\mathbf{M}\left\{\sum_{j=1}^{k_1}\frac{(n_{1j} - n_1)^2 + b}{n_{1j} + c}\right\} =$$

$$= k_1 - 1 + \frac{1}{n_1}\left[(b - c + 2)\sum_{j=1}^{k_1}\frac{1}{\pi_{ij}} - (3 - c)k_1 + 1\right] + O\left(\frac{1}{n_1^2}\right).$$

Таким образом, по крайней мере с точностью до второго порядка, π_{1j} исчезают в математическом ожидании, если $b = c - 2$. Если $b = 0$, $c = 2$, то оно равно $(k_1 - 1)\left(1 - \frac{1}{n_1}\right)$, и если $b = 1$, $c = 3$, то оно равно $(k_1 - 1) + \frac{1}{n_1}$, что для $k_1 > 2$ даже ближе к математическому ожиданию (19.91).

(Ф. Дэвид, 1950; Холдейн, 1955а.)

19.15 Для случая биномиального распределения с вероятностью успеха π показать, что МХК-оценка для π, полученная из (19.91), совпадает с МП-оценкой при любом n. Показать далее, что если число успехов не равно 0 или n, то видоизмененная МХК-оценка, найденная из (19.93), также совпадает с МП-оценкой.

19.16 Пусть y_i — пуассоновская величина с параметром θx_i, где x_i — константа, наблюдаемая вместе с y_i ($i = 1, 2, \ldots, n$). Показать, что МП-оценка для θ равна $\sum y_i / \sum x_i$ с асимптотической дисперсией $\theta / \sum x_i$ и что НК-оценка $\sum y_i x_i / \sum x_i^2$ имеет точную дисперсию $\theta \sum x_i^3 / (\sum x_i^2)^2$. Вывести отсюда, что если не все x_i равны между собой, то НК-оценка неэффективна. Объяснить этот результат.

ГЛАВА 20

ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ: ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ

20.1 В предыдущих трех главах мы имели дело с методами получения оценок для одного или нескольких неизвестных параметров; эти методы давали нам функции от выборочных значений, определявшие для любой заданной выборки единственное значение оценки. Мы, конечно, отдавали себе отчет в том, что значение оценки в каждом конкретном случае может отличаться от значения параметра и что, следовательно, при этом остается еще известная доля неопределенности. Величина этой неопределенности выражалась с помощью выборочной дисперсии статистики. Из интуитивных соображений, которыми мы пользовались до сих пор, можно было бы сказать, что, возможно, параметр θ лежит в интервале $t \pm \sqrt{\mathrm{D}(t)}$; очень возможно, что он лежит в интервале $t \pm 2\sqrt{\mathrm{D}(t)}$, и так далее. Короче говоря, что мы, по существу, могли бы сделать — это установить для θ интервал вместо отдельной точки, хотя одна точка этого интервала, а именно t, и рассматривалась бы как «наилучшая» оценка для θ.

20.2 В настоящей главе мы поближе изучим эту процедуру и взглянем на проблему оценивания с другой точки зрения. Оставим теперь попытки оценить θ с помощью функции, которая для конкретной выборки давала бы единственное значение оценки. Вместо этого мы рассмотрим, как задать промежуток, в котором лежит θ. В дальнейшем будет изучаться три метода, из которых два подобны, но не идентичны. Первый известен как метод доверительных интервалов и основан только на частотной теории вероятностей, не выражая более никакого нового принципа. Второй метод, который будем называть методом фидуциальных интервалов, требует нечто не входящее в частотную теорию. Третий основан на теореме Байеса и на одной из форм постулата Байеса (**8.4**). В настоящей главе мы постараемся изложить основные идеи и методы оценивания с помощью доверительных интервалов, принадлежащие Нейману, чей мемуар 1937 г. следует выделить особо (см. Нейман (1937b)). В 21-й главе мы будем иметь дело с соответствующими аспектами фидуциальных интервалов и байесовского оценивания.

Доверительные утверждения

20.3 Рассмотрим сначала распределение, зависящее от одного неизвестного параметра θ, и предположим, что нам дана случайная выборка $x_1, \ldots, x_n$ объема n из этого распределения. Пусть z — функция, зависящая от иксов и от θ, выборочное распределение которой не зависит от θ (примеры, приводимые ниже, покажут, что по крайней мере для некоторых случаев такую функцию можно найти). Тогда по заданной вероятности $1-\alpha$ можно найти такое значение z_1, что

$$\int_{z_1}^{\infty} dF(x) = 1 - \alpha,$$

и это выполняется для любого значения θ. В вероятностных обозначениях мы имеем тогда

$$\mathsf{P}(z \geqslant z_1) = 1 - \alpha. \tag{20.1}$$

Может случиться, что неравенство $z \geqslant z_1$ можно записать в форме $\theta \leqslant t_1$ или $\theta \geqslant t_1$, где t_1 — некоторая функция, зависящая от z_1 и иксов, но не зависящая от θ. Например, если $z = \bar{x} - \theta$, то мы имеем $\bar{x} - \theta \geqslant z_1$, и, следовательно, $\theta \leqslant \bar{x} - z_1$. Если можно переписать это неравенство таким образом, то согласно (20.1)

$$\mathsf{P}(\theta \leqslant t_1) = 1 - \alpha. \tag{20.2}$$

В более общем случае, когда распределение z может и зависеть от θ, предположим, что можно найти статистику t_1, зависящую от $1 - \alpha$ и от иксов и не зависящую от θ, такую, что (20.2) выполняется для всех θ. Тогда, пользуясь этим равенством, мы можем делать определенные утверждения о параметре θ.

20.4 Заметим прежде всего, что мы не утверждаем, что θ не превосходит константы t_1 с вероятностью $1 - \alpha$. Это утверждение (в частотной теории вероятностей) можно связывать только с изменением параметра θ в генеральной совокупности значений θ, но в общем случае мы не знаем, меняется ли θ вообще. Если параметр θ есть просто неизвестная константа, то вероятность того, что $\theta \leqslant t_1$, равна либо нулю, либо единице. Мы не знаем, какое из этих значений правильное, а знаем лишь, что одно из них верно.

По этой причине посмотрим на задачу иначе. Значение θ не случайно, но статистика t_1 является случайной величиной и изменяется от выборки к выборке. Следовательно, если мы утверждаем, что $\theta \leqslant t_1$ в каждом предложенном для решения случае, то мы в конечном счете будем правы в $(1-\alpha)$-й части

случаев. Утверждение о вероятности того, что θ меньше или равно некоторого заданного значения, не имеет смысла, кроме тривиального, упомянутого выше. Утверждение же о том, что статистика t_1 больше или равна θ (при любом θ), оказывается верным с определенной вероятностью. Следовательно, приняв за правило для любых выборочных значений утверждать, что выполнено неравенство $\theta \leqslant t_1$, мы гарантированы, что будем правы «в среднем» в $(1-\alpha)$-й части случаев.

Эта идея является основной в теории доверительных интервалов, к изложению которой мы переходим, и читатель должен убедиться в том, что он уловил ее. Подчеркнем, в частности, что доверительное утверждение выполняется, *каково бы ни было значение параметра* θ; нас интересует не повторный выбор из одной и той же совокупности, а просто повторный выбор.

20.5 Для упрощения изложения мы рассматривали только одну величину t_1 и утверждение о том, что $\theta \leqslant t_1$. На практике, однако, разыскиваются обычно две величины t_0 и t_1 такие, что для всех θ

$$\mathbf{P}(t_0 \leqslant \theta \leqslant t_1) = 1 - \alpha \text{ *)}, \qquad (20.3)$$

и делается утверждение, что θ лежит в интервале от t_0 до t_1, который называется *доверительным интервалом* для θ. Величины t_0 и t_1 называются нижней и верхней *доверительными границами* соответственно. Они зависят только от $1-\alpha$ и выборочных значений. Для любой фиксированной величины $1-\alpha$ совокупность доверительных интервалов для различных выборок определяет область, внутри которой, по утверждению, лежит значение параметра θ. Эта область называется *доверительной полосой*. Ниже будет дано графическое представление этого понятия. Величина $1-\alpha$ называется *коэффициентом доверия*.

Пример 20.1

Пусть нам дана выборка объема n из нормальной совокупности с известной дисперсией (без потери общности положим ее равной единице)

$$dF = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x-\mu)^2\right\} dx, \qquad -\infty \leqslant x \leqslant \infty.$$

Распределение выборочного среднего $\bar{x}$ есть

$$dF = \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{n}{2}(\bar{x}-\mu)^2\right\} d\bar{x}, \qquad -\infty \leqslant \bar{x} \leqslant \infty.$$

*) Почти всегда мы будем через $1-\alpha$ обозначать вероятность того, что интервал накрывает неизвестный параметр. В литературе нет единой практики и часто пишут α вместо $1-\alpha$. Наше соглашение является в настоящее время более распространенным.

Из таблиц нормального интеграла известно, что вероятность положительного отклонения от среднего не более чем на два стандартных отклонения равна 0,97725. Следовательно, имеем

$$\mathbf{P}(\bar{x} - \mu \leqslant 2/\sqrt{n}) = 0{,}97725,$$

что эквивалентно соотношению

$$\mathbf{P}(\bar{x} - 2/\sqrt{n} \leqslant \mu) = 0{,}97725.$$

Таким образом, если мы утверждаем, что μ больше или равно $\bar{x} - 2/\sqrt{n}$, то мы будем правы в среднем примерно в 97,725% случаев.

Аналогично имеем

$$\mathbf{P}(\bar{x} - \mu \geqslant -2/\sqrt{n}) = \mathbf{P}(\mu \leqslant \bar{x} + 2/\sqrt{n}) = 0{,}97725.$$

Отсюда, объединяя оба результата, получаем

$$\mathbf{P}(\bar{x} - 2/\sqrt{n} \leqslant \mu \leqslant \bar{x} + 2/\sqrt{n}) = 1 - 2\,(1 - 0{,}97725) = 0{,}9545. \tag{20.4}$$

Следовательно, если мы утверждаем, что μ лежит в интервале $\bar{x} \pm 2/\sqrt{n}$, то будем правы в среднем примерно в 95,45% случаев.

Обратно, если задан коэффициент доверия, то мы можем легко найти с помощью таблиц нормального интеграла отклонение d такое, что

$$\mathbf{P}(\bar{x} - d/\sqrt{n} \leqslant \mu \leqslant \bar{x} + d/\sqrt{n}) = 1 - \alpha.$$

Например, если $1 - \alpha = 0{,}8$, то $d = 1{,}28$. Утверждая в этом случае, что μ лежит в интервале $\bar{x} \pm 1{,}28/\sqrt{n}$, мы имеем 4 шанса против 1 быть правыми.

Читатель, для которого этот подход нов, возможно спросит, не есть ли это окольный метод использования стандартной ошибки для установления пределов изменения оценки среднего. В некотором отношении это так. Действительно, приведенный пример показывает, что использование стандартной ошибки для среднего в нормальной выборке может быть логически оправдано без привлечения новых методов, отличных от уже имеющихся в теории вероятностей. В частности, мы не использовали постулат Байеса (8.4).

Отметим интересный момент в нашем примере, заключающийся в том, что нижняя и верхняя доверительные границы, полученные выше, равно отстоят от среднего $\bar{x}$. Это не является необходимым, и легко видеть, что мы можем получить любое число альтернативных границ для того же коэффициента доверия $1 - \alpha$. Предположим, например, что $1 - \alpha = 0{,}9545$, и выбе-

рем два числа α_0 и α_1, подчиненных условию

$$\alpha_0 + \alpha_1 = \alpha = 0{,}0455,$$

скажем, $\alpha_0 = 0{,}01$ и $\alpha_1 = 0{,}0355$. Из таблиц нормального интеграла имеем

$$\mathsf{P}(\bar{x} - \mu \leqslant 2{,}326/\sqrt{n}) = 0{,}99,$$
$$\mathsf{P}(\bar{x} - \mu \geqslant -1{,}806/\sqrt{n}) = 0{,}9645,$$

и, следовательно,

$$\mathsf{P}\left(\bar{x} - \frac{2{,}326}{\sqrt{n}} \leqslant \mu \leqslant \bar{x} + \frac{1{,}806}{\sqrt{n}}\right) = 0{,}9545. \qquad (20.5)$$

Таким образом, с тем же коэффициентом доверия можно утверждать, что μ лежит в интервале от $\bar{x} - 2/\sqrt{n}$ до $\bar{x} + 2/\sqrt{n}$ или в интервале от $\bar{x} - 2{,}326/\sqrt{n}$ до $\bar{x} + 1{,}806/\sqrt{n}$. В каждом случае мы будем правы в среднем примерно в 95,45% случаев.

Заметим, что в первом случае интервал имеет длину $4/\sqrt{n}$, в то время как во втором случае его длина равна $4{,}132/\sqrt{n}$. При прочих равных условиях нам следует выбрать первые границы, так как они заключают параметр в более узкий интервал. Ниже мы рассмотрим этот вопрос более детально. Не всегда случается, что имеется бесконечное число возможных доверительных интервалов или что выбор между ними может быть сделан на столь же понятных основаниях, как в этом примере.

Графическое представление

20.6 В ряде простых случаев, включая пример 20.1, доверительные границы могут быть представлены в удобной графической форме. Возьмем две ортогональные оси: OX, относящуюся к наблюденному значению $\bar{x}$, и OY, отвечающую μ (рис. 20.1).

Две изображенные прямые линии описываются уравнениями

$$\mu = \bar{x} + 2, \quad \mu = \bar{x} - 2.$$

Следовательно, для любой точки между линиями имеем

$$\bar{x} - 2 \leqslant \mu \leqslant \bar{x} + 2.$$

Поэтому, если для любого наблюденного значения $\bar{x}$ мы найдем соответствующие ему две ординаты на прямых, то получим две доверительных границы. Вертикальный интервал между границами есть доверительный интервал (показан на рисунке для $\bar{x} = 1$), а вся зона между прямыми образует доверительную полосу. Мы можем называть эти две прямые нижней и верхней доверительными линиями соответственно.

Графическое изображение сделано для случая $n = 1$ примера 20.1. При разных значениях n будут получаться различные

линии, все параллельные прямой $\mu = \bar{x}$ и приближающиеся одна к другой с ростом n. Для всевозможных n их можно нанести на один рисунок, и полученная диаграмма оказывается полезной в практической работе для нахождения доверительных границ.

С другой стороны, можно изменять коэффициент доверия $1 - \alpha$ (в нашем примере он равен 0,9545). Тогда на одной и той же диаграмме, относящейся к некоторому фиксированному значению n, мы можем нанести ряд пар доверительных линий, каждая из которых соответствует некоторому выбранному значению $1 - \alpha$. В этом случае, конечно, линии удаляются одна от другой с ростом $1 - \alpha$. Во многих практических ситуациях нас интересует изменение доверительного интервала в зависимости от $1 - \alpha$, и мы можем одновременно для некоторого набора значений α сделать утверждения вида (20.3), каждое из которых в среднем будет справедливо в соответствующей доле случаев. Эта процедура может быть доведена до своей крайней формы, когда одновременно рассматривается вся совокупность значений $1 - \alpha$ в промежутке $(0, 1)$. Так порождается «доверительное распределение» параметра — термин, принадлежащий Д. Коксу (см., например, 1958b): одновременно мы имеем бесконечную последовательность доверительных утверждений, причем с ростом величины $1 - \alpha$ каждое последующее утверждение содержится в предыдущем.

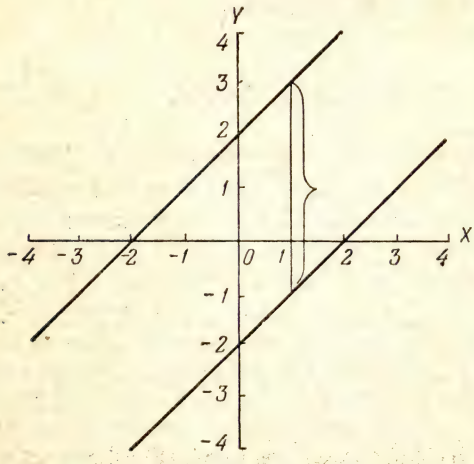


Рис. 20.1. Доверительные границы в примере 20.1 для $n = 1$.

Центральные и нецентральные интервалы

20.7 В примере 20.1 выборочное распределение, на котором основывались доверительные интервалы, было симметричным, поэтому, беря одинаковые отклонения от среднего, мы получали одинаковые значения для

$$1 - \alpha_0 = \mathsf{P}(t_0 \leqslant \theta) \quad \text{и} \quad 1 - \alpha_1 = \mathsf{P}(\theta \leqslant t_1).$$

В общем случае, беря одинаковые отклонения, такого результата получить нельзя, но α_0 и α_1 можно выбирать произвольно при условии, что $\alpha_0 + \alpha_1 = \alpha$.

Если α_0 и α_1 взяты одинаковыми, то интервал будем называть *центральным*. Тогда

$$\mathsf{P}(t_0 > \theta) = \mathsf{P}(\theta > t_1) = \alpha/2. \qquad (20.6)$$

В противном случае интервалы будем называть *нецентральными*. Следует заметить, что центральность в этом смысле не означает, что доверительные границы равно отстоят от выборочной статистики, если только ее выборочное распределение не симметрично *).

20.8 При отсутствии других соображений обычно используют центральные интервалы, но при некоторых обстоятельствах нецентральные интервалы оказываются более подходящими. Предположим, например, что в медицинском препарате оценивается содержание вещества, которое в больших дозах ядовито. Тогда ошибаться мы можем только в безопасную сторону, а в данном случае более опасным является превышение истинного значения над оцененным. Поэтому было бы желательно взять α_1 равным нулю:

$$\mathsf{P}(\theta \leqslant t_1) = 1, \quad \mathsf{P}(t_0 \leqslant \theta) = 1 - \alpha,$$

для того, чтобы быть уверенным в том, что θ не превосходит t_1. Но если наша статистика распределена на бесконечном интервале, то $\mathsf{P}(\theta \leqslant t_1) = 1$ только при t_1, равном бесконечности, и поэтому мы должны довольствоваться величиной α_1, очень близкой к нулю.

Если же, например, оценивается доля жизнеспособных семян в выборке из материала, поступающего на продажу, то важнее точность нижней границы, чем верхней: с точки зрения фермера недостаток всхожести более серьезен, чем избыток. При таких обстоятельствах α_0 следует взять по возможности малым для того, чтобы быть как можно более уверенным относительно наименьшего значения доли жизнеспособных семян. Ситуация такого рода часто возникает при установлении качества изготовленного продукта; продавец желает гарантировать минимальный стандарт и значительно менее озабочен тем, чтобы его продукт превосходил ожидание.

20.9 Следует заметить, что при определенных обстоятельствах достаточно знать, что $\mathsf{P}(t_0 \leqslant \theta \leqslant t_1) \geqslant 1 - \alpha$. Тогда, утверждая, что θ лежит в интервале от t_0 до t_1, мы не ошибемся по крайней мере в $(1-\alpha)$-й части случаев. Например, математические

*) На самом деле из симметрии распределения выборочной статистики не следует, что доверительные границы будут равноотстоящими от нее. Пусть, например, оценивается параметр θ в нормальном распределении со средним θ и дисперсией θ^2. Тогда, пользуясь тем, что $z = (\bar{x} - \theta)/(\theta/\sqrt{n})$ имеет нормальное $(0, 1)$ распределение, получаем для θ доверительные границы вида $\bar{x}/(1 \pm d/\sqrt{n})$, не равноотстоящие от $\bar{x}$. (*Прим. ред.*)

трудности при установлении точных доверительных границ для заданного $1-\alpha$ или теоретические трудности при дискретном распределении могут заставить нас довольствоваться этим неравенством вместо равенства (20.3).

Пример 20.2

Найти доверительный интервал для вероятности «успеха» ω при выборке по качественному признаку.

В выборке объема n распределение числа успехов определяется членами биномиального разложения $(\chi+\omega)^n$, где $\chi = 1-\omega$. Найдем доверительные границы для случая $n=20$ и при коэффициенте доверия 0,95.

Прежде всего нам нужна биномиальная функция распределения. Данная ниже таблица содержит ее значения для определенных ω вплоть до $\omega=0{,}5$ (для бóльших ω распределения получаются из симметрии). Для точного построения доверительной полосы нам нужна более детальная информация, которую можно получить из подробных таблиц биномиального распределения, указанных в **5.7**. Приводимая же здесь таблица будет служить лишь для иллюстрации.

Окончательные цифры могут иметь ошибку в одну-две единицы последнего знака вследствие ошибок округления, но при степени точности, рассматриваемой здесь, это не должно нас беспокоить.

Доля успехов p	$\omega=0{,}1$	$\omega=0{,}2$	$\omega=0{,}3$	$\omega=0{,}4$	$\omega=0{,}5$
0,00	0,1216	0,0115	0,0008	—	—
0,05	0,3918	0,0691	0,0076	0,0005	—
0,10	0,6770	0,2060	0,0354	0,0036	0,0002
0,15	0,8671	0,4114	0,1070	0,0159	0,0013
0,20	0,9569	0,6296	0,2374	0,0509	0,0059
0,25	0,9888	0,8042	0,4163	0,1255	0,0207
0,30	0,9977	0,9133	0,6079	0,2499	0,0577
0,35	0,9997	0,9678	0,7722	0,4158	0,1316
0,40	1,0001	0,9900	0,8866	0,5955	0,2517
0,45	1,0002	0,9974	0,9520	0,7552	0,4119
0,50	—	0,9994	0,9828	0,8723	0,5881
0,55	—	0,9999	0,9948	0,9433	0,7483
0,60	—	1,0000	0,9987	0,9788	0,8684
0,65	—	—	0,9997	0,9934	0,9423
0,70	—	—	0,9999	0,9983	0,9793
0,75	—	—	—	0,9996	0,9941
0,80	—	—	—	0,9999	0,9987
0,85	—	—	—	—	0,9998
0,90	—	—	—	—	1,0000
0,95	—	—	—	—	—

Заметим вначале, что варианта p дискретна. С другой стороны, мы собираемся рассматривать любые значения ω в интервале от 0 до 1. При заданном ω нельзя в общем случае найти границы для p такие, чтобы $1-\alpha$ было равно в точности 0,95, но мы будем брать в качестве p относительную частоту, которая дает коэффициент доверия, не меньший 0,95. Будем рассматривать только центральные интервалы. Таким образом, для заданного ω нужно найти ω_0 и ω_1 такие, чтобы

$$\mathbf{P}(p \geqslant \omega_0) \geqslant 0{,}975,$$
$$\mathbf{P}(p \leqslant \omega_1) \geqslant 0{,}975$$

и чтобы неравенства для вероятности $\mathbf{P}$ были как можно ближе к равенству.

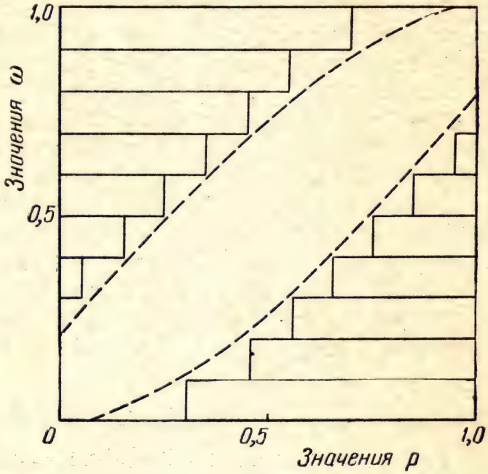


Рис. 20.2. Доверительные границы для биномиального параметра.

Рассмотрим диаграмму, представленную на рис. 20.2. Для любого заданного ω из таблицы можно найти значения ω и ω_1 такие, чтобы $\mathbf{P}(p \geqslant \omega_0) \geqslant 0{,}975$ и $\mathbf{P}(p \leqslant \omega_1) \geqslant 0{,}975$. Заметим, что при определении ω_1 функция распределения дает вероятность получения доли успехов, меньшей или равной p. Следовательно, дополнение функции распределения до единицы дает вероятность получения доли успехов, строго большей, чем p. Например, на горизонтали, проходящей через $\omega = 0{,}1$, мы имеем $\omega_0 = 0$ и $\omega_1 = 0{,}25$, найденные из нашей таблицы; при $\omega = 0{,}4$ имеем $\omega_0 = 0{,}15$ и $\omega_1 = 0{,}60$. Полученные точки лежат на ступенчатых линиях. Например, когда $\omega = 0{,}3$, наибольшее значение ω_0 такое, что $\mathbf{P}(p \geqslant \omega_0) \geqslant 0{,}975$, есть 0,1. С возрастанием ω до 0,4 величина ω_0 возрастает до 0,20. Где-то между 0,3 и 0,4 находится значение ω такое, что $\mathbf{P}(p \geqslant 0{,}1)$ равна в точности 0,975. Если бы мы табулировали вероятности при более мелких интервалах между значениями ω, то ступенчатые кривые несколько изменились бы, и в пределе, если вычислить значения ω такие, что в точности $\mathbf{P}(p \geqslant \omega_0) = 0{,}975$, мы получили бы точки, лежащие внутри приведенных ступенчатых кривых. Эти точки на рис. 20.2 соединены пунктирными линиями.

Зона между ступенчатыми линиями является доверительной полосой. Для любого p вероятность ошибиться, утверждая, что ω

лежит внутри полосы, не превосходит 0,05. Через наблюденное значение p на оси абсцисс проведем вертикаль, точки ее пересечения с соответствующими линиями, дающими ω_0, ω_1, определяют величины p_0 и p_1. Чуть ниже мы покажем, что они представляют собой искомые границы.

Ниже (**20.22**) мы рассмотрим более сложный метод обращения с дискретностью.

Стоит заметить, что точки на кривых рис. 20.2 строились следующим образом: для выбранной ординаты ω находились соответствующие абсциссы ω_0 и ω_1. Диаграмма строилась, так сказать, *по горизонтали*. Однако при ее применении она читается *по вертикали*, т. е. для наблюденной абсциссы p мы считываем с нее две величины p_0 и p_1 и утверждаем, что $p_0 \leqslant \omega \leqslant p_1$. Поучительно пронаблюдать, как это изменение точки зрения может быть оправдано без обращения к постулату Байеса.

Рассматривая диаграмму по горизонтали, мы видим, что для любого заданного ω наблюдение попадает в доверительную полосу с вероятностью $\geqslant 1-\alpha$. Тогда и только тогда, когда наблюдение находится внутри полосы, истинное значение ω будет содержаться между p_0 и p_1. Следовательно, последнее событие имеет вероятность $\geqslant 1-\alpha$, каково бы ни было истинное значение ω.

Доверительные интервалы для больших выборок

20.10 Мы видели в **18.16**, что первая производная логарифма функции правдоподобия, при условиях регулярности, распределена асимптотически нормально со средним, равным нулю, и

$$\mathrm{D}\left(\frac{\partial \log L}{\partial \theta}\right) = \mathrm{M}\left\{\left(\frac{\partial \log L}{\partial \theta}\right)^2\right\} = -\mathrm{M}\left\{\frac{\partial^2 \log L}{\partial \theta^2}\right\}. \qquad (20.7)$$

Используя этот факт, можно находить доверительные интервалы для θ в случае больших выборок. Обозначим

$$\psi = \frac{\partial \log L}{\partial \theta} \Big/ \left[\mathrm{M}\left\{\left(\frac{\partial \log L}{\partial \theta}\right)^2\right\}\right]^{1/2}, \qquad (20.8)$$

так что ψ является, при больших выборках, стандартной нормальной случайной величиной. Если ψ — монотонная функция от θ, так что неравенство для одной величины можно преобразовать в неравенство для другой, то из нормального интеграла можно определить доверительные границы для θ. Следующий пример иллюстрирует эту процедуру.

Пример 20.3

Рассмотрим вновь задачу из примера 20.1. Мы видели в примере 17.6, что для этого случая

$$\frac{\partial \log L}{\partial \mu} = n(\bar{x} - \mu), \qquad (20.9)$$

откуда

$$-\frac{\partial^2 \log L}{\partial \mu^2} = n. \tag{20.10}$$

Из (20.7) и (20.8) находим, что при больших n величина

$$\psi = (\bar{x} - \mu)\sqrt{n} \tag{20.11}$$

имеет нормальное распределение с единичной дисперсией (конечно, в нашем частном случае это верно для любых n). Доверительные границы для μ могут быть затем найдены так же, как в примере 20.1.

Пример 20.4

Рассмотрим распределение Пуассона, общий член которого есть

$$f(x, \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, \ldots$$

В примере 17.8 мы видели, что

$$\frac{\partial \log L}{\partial \lambda} = \frac{n}{\lambda}(\bar{x} - \lambda). \tag{20.12}$$

Следовательно,

$$-\frac{\partial^2 \log L}{\partial \lambda^2} = \frac{n\bar{x}}{\lambda^2}$$

и

$$\mathbf{M}\left(-\frac{\partial^2 \log L}{\partial \lambda^2}\right) = \frac{n}{\lambda}. \tag{20.13}$$

Таким образом, из (20.7) и (20.8) имеем

$$\psi = (\bar{x} - \lambda)\sqrt{n/\lambda}. \tag{20.14}$$

Например, при $1 - \alpha = 0{,}95$, что соответствует нормальному уклонению $\pm 1{,}96$, для центральных доверительных границ находим

$$(\bar{x} - \lambda)\sqrt{n/\lambda} = \pm 1{,}96.$$

Отсюда для нахождения λ получаем уравнение

$$\lambda^2 - \left(2\bar{x} + \frac{3{,}84}{n}\right)\lambda + \bar{x}^2 = 0,$$

из которого

$$\lambda = \bar{x} + \frac{1{,}92}{n} \pm \sqrt{\frac{3{,}84\bar{x}}{n} + \frac{3{,}69}{n^2}},$$

причем знаки плюс и минус перед квадратным корнем дают верхнюю и нижнюю границы соответственно.

С точностью до членов порядка $n^{-1/2}$ это решение эквивалентно

$$\lambda = \bar{x} \pm 1{,}96\sqrt{\bar{x}/n}; \tag{20.15}$$

в этом случае, как и следовало ожидать, верхняя и нижняя границы равно отстоят от среднего $\bar{x}$.

20.11 Процедура, использованная для получения (20.15), требует некоторых дальнейших исследований. Если мы имеем вероятности типа

$$\mathbf{P}(\theta \leqslant t) \quad \text{или} \quad \mathbf{P}(\theta^{-1} \leqslant t), \qquad \theta > 0,$$

то их можно немедленно «обратить» и получить

$$\mathbf{P}(t \geqslant \theta) \quad \text{или} \quad \mathbf{P}(t^{-1} \leqslant \theta).$$

Но может встретиться более сложная форма зависимости, такая, как

$$\mathbf{P}\{g(t, \theta) \leqslant 0\},$$

где g есть, например, полином относительно t либо θ либо сразу относительно обеих переменных, степени выше первой. Задача преобразования такого выражения в выражение, имеющее вид интервала для θ, может быть далеко не простой.

Рассмотрим еще раз (20.14) в форме

$$\psi = n^{1/2}(\bar{x} - \lambda)/\lambda^{1/2}. \tag{20.16}$$

Возьмем коэффициент доверия $1-\alpha$, и пусть соответствующие значения ψ равны ψ_0 и ψ_1, т. е.

$$\mathbf{P}\{\psi_0 \leqslant \psi \leqslant \psi_1\} = 1 - \alpha. \tag{20.17}$$

Уравнение (20.16) может быть переписано так:

$$\lambda^2 - (2\bar{x} + n^{-1}\psi^2)\lambda + \bar{x}^2 = 0, \tag{20.18}$$

и если интервалы для ψ являются центральными, т. е. $\psi_0 = -\psi_1$, то уравнение (20.18) будет иметь одни и те же корни при $\psi = \psi_0$ и при $\psi = \psi_1$. Кроме того, корни всегда действительны. Пусть λ_0, λ_1 — корни уравнения с $\psi = \psi_0$ (или ψ_1), и пусть λ_1 — больший. Тогда при изменении ψ от $-\infty$ до ψ_0, как видно из (20.18), λ изменяется от $+\infty$ до λ_1; когда ψ пробегает значения от ψ_0 до ψ_1, то λ (убывая) пробегает значения от λ_1 до λ_0, и когда ψ изменяется от ψ_1 до $+\infty$, λ изменяется от λ_0 до $-\infty$. Таким образом, соотношение

$$\mathbf{P}(\psi_0 \leqslant \psi \leqslant \psi_1) = 1 - \alpha$$

эквивалентно соотношению

$$\mathbf{P}(\lambda_0 \leqslant \lambda \leqslant \lambda_1) = 1 - \alpha,$$

и наши доверительные интервалы имеют требуемый вид.

Полезно рассмотреть это графически, как показано на рис. 20.3.

Из (20.16) видно, что при заданных n и ψ кривая, связывающая λ (ординату) и $\bar{x}$ (абсциссу), может быть представлена в виде

$$(\lambda - \bar{x})^2 = k\lambda, \qquad (20.19)$$

где k — положительная постоянная. При различных k это параболы, проходящие через начало координат и имеющие главную ось $\lambda = \bar{x}$. Линия $\lambda = \bar{x}$ соответствует $k = 0$ или $n = \infty$, а две другие кривые (без масштаба) приведены для значений n_1 и n_2 ($n_1 < n_2$). Из проведенных выше рассуждений следует, что при заданном n значения λ, соответствующие значениям ψ, лежащим *внутри* интервала от ψ_0 до ψ_1, лежат *внутри* соответствующей параболы. Очевидно также, что парабола для n_1 лежит целиком внутри параболы для любого меньшего n.

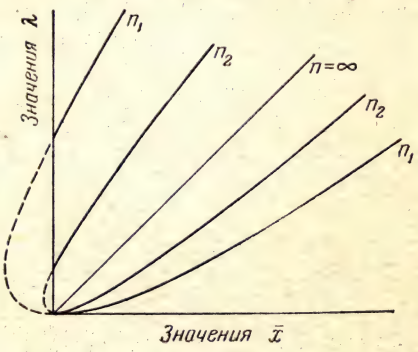


Рис. 20.3. Доверительные параболы (20.19) для различных k или n.

Итак, при любом данном $\bar{x}$ можно определить ординаты для двух соответствующих значений λ и утверждать, что неизвестный параметр λ лежит между ними. Доверительные линии на рис. 20.3 являются выпуклыми и вложенными, как и доверительные линии для биномиального распределения в примере 20.2.

20.12 Рассмотрим теперь более сложный случай. Предположим, что мы имеем статистику t, с помощью которой можем найти границы t_0 и t_1, не зависящие от θ, при заданном коэффициенте доверия $1 - \alpha$. И пусть

$$t = a\theta^3 + b\theta^2 + c\theta + d, \qquad (20.20)$$

где a, b, c, d — константы *). Иногда, но не всегда, будут существовать три действительных значения θ, соответствующих одному значению t. Как ими пользоваться, чтобы определить границы для θ?

Обратимся к графическому представлению. На рис. 20.4 (вновь без масштаба) мы считаем θ ординатой, а t — абсциссой.

*) Формулировка непонятна: например, в левой части (20.20) стоит статистика, а в правой — функция от параметра. По-видимому, имеется в виду, что распределение разности $\psi = a\theta^3 + b\theta^2 + c\theta + d - t$ не зависит от θ. Тогда соотношение типа (20.17) приводит к неравенству $t_0 \leqslant a\theta^3 + b\theta^2 + c\theta + d \leqslant t_1$, где $t_0 = \psi_0 + t$ и $t_1 = \psi_1 + t$, и дальнейшее изложение посвящено обращению этого неравенства. (*Прим. ред.*)

Пусть константы таковы, что кривая третьего порядка имеет действительные максимум и минимум, как это показано на рисунке. Для различных значений t кривая (20.20) перемещается вдоль оси t. Чтобы избежать путаницы, предположим, что приведены линии только для одного значения n. Пусть также $a > 0$.

При данном значении t, например t_0, получим кривую третьего порядка, как показано на рисунке, такую, что в области справа от нее $a\theta^3 + b\theta^2 + c\theta + d > t_0$, а слева $< t_0$. Аналогично при t_1. С соответствующим коэффициентом доверия можно теперь сказать, что для наблюденного t границы для θ получаются путем чтения по вертикали вдоль оси ординат в точке t.

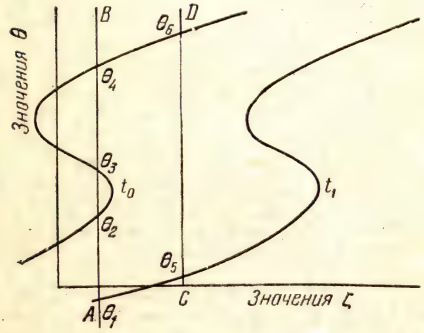


Рис. 20.4. Доверительные кривые третьего порядка (см. текст).

Здесь начинаются трудности. Если мы берем такое значение, как на рисунке вдоль AB, то должны утверждать, что θ лежит в разорванном интервале $\theta_1 \leqslant \theta \leqslant \theta_2$ и $\theta_3 \leqslant \leqslant \theta \leqslant \theta_4$. С другой стороны, на CD получается односвязный промежуток $\theta_5 \leqslant \theta \leqslant \theta_6$.

Для любителя патологической математики не составит труда построение дальнейших примеров, в которых интервалы разбиты на еще большее число частей или в которых мы должны утверждать, что параметр θ лежит *вне* замкнутого интервала. (Некоторые важные примеры см. в работах Филлера (1954) и С. Дэвида (1954).) См. также упражнение 28.21 ниже.

20.13 Следует заметить, что в таких случаях утверждения относительно интервалов могут быть сделаны с прежней точностью. Вопрос в том, есть ли от них польза и дают ли они решение задачи, с которой мы начали, состоящей в том, чтобы установить окрестность для значения параметра. Можно ли считать их доверительными интервалами или они уже не могут так называться?

Нельзя дать простого ответа на подобные вопросы, но мы постараемся изложить наше собственное мнение на эту тему.

(а) Наиболее удовлетворительна та ситуация, при которой доверительные линии монотонны, т. е. любая прямая, параллельная оси ординат, пересекает каждую линию только в одной точке. Тогда мы получаем утверждение, что параметр лежит внутри односвязного интервала. Далее желательно, чтобы при фиксированном α доверительная полоса для любого n лежала внутри полосы для меньшего n и чтобы при фиксированном n полоса

для любого $1-\alpha$ лежала внутри полосы для большего $1-\alpha$. Эти условия выполнены в примерах 20.1—20.4.

(б) Случай, где эти условия не выполняются, должен быть рассмотрен по существу дела. Могут быть примеры, в которых появляются неодносвязные интервалы, такие, как на рис. 20.4, и они оказываются приемлемыми. Там, где это возможно, доверительные области должны быть представлены графически. Нужно избегать механического «обращения» вероятностных утверждений, не принимающего во внимание эти моменты.

20.14 Мы можем на данном этапе отметить трудности другого сорта. Когда рассматривалось квадратное уравнение (20.18), то было отмечено, что в условиях задачи корни всегда действительные. Однако может случиться, что для некоторых коэффициентов доверия невозможно построить действительные доверительные интервалы. Продемонстрируем это на следующем примере.

Пример 20.5

Если $x_1, \ldots, x_n$ — выборка объема n из нормальной совокупности с единичной дисперсией и средним μ, то статистика $\chi^2 = \sum (x-\mu)^2$ имеет распределение хи-квадрат с n степенями свободы. Для выбранного доверительного коэффициента $1-\alpha$ можно определить χ_0^2 и χ_1^2 (для упрощения изложения можно строить центральный интервал) такие, что

$$\mathsf{P}\{\chi_0^2 \leqslant \chi^2 \leqslant \chi_1^2\} = 1-\alpha. \tag{20.21}$$

Если $s^2 = \sum (x-\bar{x})^2/n$, то имеем тождество $\chi^2 = \sum (x-\mu)^2 = n\{s^2 + (\bar{x}-\mu)^2\}$, и, следовательно, границы для $(\bar{x}-\mu)^2$ имеют вид

$$\frac{\chi_0^2}{n} - s^2 \leqslant (\bar{x}-\mu)^2 \leqslant \frac{\chi_1^2}{n} - s^2. \tag{20.22}$$

Может случиться, что s^2 больше, чем χ_1^2/n. Тогда (поскольку $\chi_0^2 < \chi_1^2$) неравенство (20.22) утверждает, что $(\bar{x}-\mu)^2$ лежит между двумя отрицательными величинами. Что можно сделать с таким утверждением?

Дело станет яснее, если вновь обратиться к геометрической интерпретации. Так как χ^2 теперь зависит от двух статистик, s и $\bar{x}$ (которые независимы), то нам для изображения нужны три измерения, соответствующие μ, s и $\bar{x}$. Рис. 20.5 дает геометрическое представление.

Величина χ^2 постоянна на поверхностях $(\bar{x}-\mu)^2 + s^2 = \text{const}$. Для фиксированного μ (т. е. в плоскости, перпендикулярной оси μ) эти поверхности пересекают плоскость $\mu = \text{const}$ по окружностям с центром в точке $(\mu, 0)$. Все эти центры лежат в пло-

скости $(\mu, \bar{x})$ на прямой с уравнением $\mu = \bar{x}$, а поверхности с постоянной величиной χ^2 являются цилиндрами, имеющими эту прямую своей осью (они не являются круговыми цилиндрами; только сечения, перпендикулярные к оси μ, являются кругами).

Кроме того, цилиндр для χ_1^2 полностью заключает в себе цилиндр для χ_0^2, как это представлено на рисунке. При данных наблюденных значениях $\bar{x}$, s проведем прямую, параллельную оси μ. Если она пересекает каждый цилиндр в двух точках: μ_{00}, μ_{01} для χ_0^2 и μ_{10}, μ_{11} для χ_1^2, то мы утверждаем, что $\mu_{00} \leqslant \mu \leqslant \mu_{10}$ и $\mu_{01} \leqslant \mu \leqslant \mu_{11}$ (эти два интервала соответствуют двум знакам при извлечении квадратного корня в (20.22)).

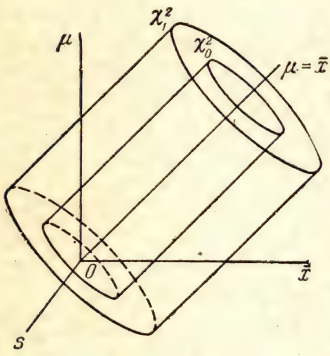


Рис. 20.5. Доверительные цилиндры (20.22) (см. текст).

Существенным моментом в настоящем примере является то, что проведенная прямая может вообще не пересекать цилиндров. Тогда корни для μ в (20.22) будут мнимыми. Такая ситуация не может возникнуть, к примеру, в биномиальном случае примера 20.2, где каждая прямая, параллельная оси ω в интервале $0 \leqslant p \leqslant 1$, должна пересечь доверительную полосу. Так обычно бывает, когда параметр θ имеет единственную достаточную статистику t, с помощью которой строятся интервалы (хотя при этом возможны трудности с обращением неравенств типа рассмотренных в **20.11—13**). Но этого уже может не произойти, когда мы используем более чем одну статистику и имеем дело более чем с двумя измерениями, как это было в рассмотренном примере.

В таких случаях, как нам кажется, мы должны признать, что перед нами неразрешимая задача*). *Используя эти конкретные статистики*, мы хотим высказать с коэффициентом доверия $1 - \alpha$ утверждение, которое имело бы силу для всех наблюденных $\bar{x}$ и s. Этого сделать нельзя. Это возможно только для определенных множеств значений $\bar{x}$ и s, а именно для таких, при которых границы в (20.22) положительны. При определенных значениях $\bar{x}$ и s мы можем снизить наш доверительный уровень, увеличить радиусы цилиндров и обеспечить, чтобы линия, проходящая через $\bar{x}$, s, пересекала цилиндры. Но, каким бы низким мы ни

*) На наш взгляд, Нейман сказал бы, что такие интервалы не являются доверительными интервалами в его смысле (нарушены условия **20.28**, приводимые ниже). Другие авторы использовали выражения для интервалов, полученные обращением вероятностного утверждения, без учета этих условий.

брали уровень, всегда может появиться выборка, для которой таким путем нельзя установить границы для μ.

В настоящем примере ясно, как поправить дело. Мы выбрали неверный метод построения доверительных интервалов; в самом деле, если воспользоваться методом из примера 20.1 и установить границы для $\bar{x} - \mu$ с помощью нормального распределения, то трудности не возникнет. В этом случае $\bar{x}$ есть достаточная статистика для μ. В общем случае, когда не существует одномерной достаточной статистики, может оказаться, что трудности избежать нельзя, и ее нужно принять открыто, если и не так, как мы предлагаем, то любым другим столь же явным образом.

20.15 Мы возвращаемся к аппроксимации доверительных интервалов (для большой выборки), которую обсуждали в **20.10**. Если мы получаем недостаточную точность приближения, считая, что ψ распределена нормально для выборки объема n, с которой мы имеем дело, то может быть получена более точная аппроксимация. Для этого найдем высшие моменты величины ψ и воспользуемся разложением типа Корниша — Фишера (**6.25—6**). Используя (17.19), запишем

$$I = \mathbf{M}\left(\frac{\partial \log L}{\partial \theta}\right)^2 = -\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right), \tag{20.23}$$

$$J = \frac{\partial \log L}{\partial \theta}. \tag{20.24}$$

Из (17.18) при условиях регулярности получаем

$$\varkappa_1(J) = 0, \tag{20.25}$$

откуда

$$\varkappa_2(J) = I. \tag{20.26}$$

Теперь докажем, что

$$\varkappa_3(J) = 3\frac{\partial I}{\partial \theta} + 2\mathbf{M}\left(\frac{\partial^3 \log L}{\partial \theta^3}\right), \tag{20.27}$$

$$\varkappa_4(J) = 6\frac{\partial^2 I}{\partial \theta^2} + 8\frac{\partial}{\partial \theta}\mathbf{M}\left(\frac{\partial^3 \log L}{\partial \theta^3}\right) - 3\mathbf{M}\left(\frac{\partial^4 \log L}{\partial \theta^4}\right) + 3\mathbf{D}\left(\frac{\partial^2 L}{\partial \theta^2}\right). \tag{20.28}$$

Действительно, дифференцируя I, имеем

$$\frac{\partial I}{\partial \theta} = 2\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\,\frac{\partial \log L}{\partial \theta}\right) + \mathbf{M}\left(\frac{\partial \log L}{\partial \theta}\right)^3, \tag{20.29}$$

а дифференцируя равенство

$$0 = I + \mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right),$$

получаем

$$0 = \frac{\partial I}{\partial \theta} + \mathbf{M}\left(\frac{\partial^3 \log L}{\partial \theta^3}\right) + \mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\,\frac{\partial \log L}{\partial \theta}\right). \tag{20.30}$$

Исключая $\mathbf{M}\{(\partial^2 \log L/\partial\theta^2)(\partial \log L/\partial\theta)\}$ из (20.29) и (20.30), приходим к (20.27).

Дифференцируя дважды оба соотношения для I, данные в (20.23), и исключая $\mathbf{M}\{(\partial^2 \log L/\partial\theta^2)(\partial \log L/\partial\theta)^2\}$, находим

$$6\frac{\partial^2 I}{\partial\theta^2} = \mathbf{M}\left(\frac{\partial \log L}{\partial\theta}\right)^4 - 8\mathbf{M}\left(\frac{\partial^3 \log L}{\partial\theta^3}\frac{\partial \log L}{\partial\theta}\right) - \\ - 5\mathbf{M}\left(\frac{\partial^4 \log L}{\partial\theta^4}\right) - 3\mathbf{M}\left(\frac{\partial^2 \log L}{\partial\theta^2}\right)^2.$$

Используя соотношение

$$\frac{\partial}{\partial\theta}\mathbf{M}\left(\frac{\partial^3 \log L}{\partial\theta^3}\right) = \mathbf{M}\left(\frac{\partial^3 \log L}{\partial\theta^3}\frac{\partial \log L}{\partial\theta}\right) + \mathbf{M}\left(\frac{\partial^4 \log L}{\partial\theta^4}\right)$$

и переходя к семиинвариантам, приходим к (20.28). Эти формулы принадлежат Бартлетту (1953).

Беря первые четыре члена в (6.54) с $l_1 = l_2 = 0$, получаем статистику

$$T(\theta) = \frac{1}{\sqrt{I}}\left[\frac{\partial \log L}{\partial\theta} - \frac{1}{6}\frac{\varkappa_3(J)}{I^2}\left\{\left(\frac{\partial \log L}{\partial\theta}\right)^2 - I\right\} - \right. \\ \left. - \frac{1}{24}\frac{\varkappa_4(J)}{I^3}\left\{\left(\frac{\partial \log L}{\partial\theta}\right)^3 - 3I\frac{\partial \log L}{\partial\theta}\right\}\right], \quad (20.31)$$

которая распределена (в следующем порядке аппроксимации) нормально со средним нуль и единичной дисперсией. Первый член в этой формуле мы обозначили ψ. Поправочный член включает в себя нормированные семиинварианты величины J или, что эквивалентно, семиинварианты величины ψ.

Пример 20.6

Рассмотрим задачу построения доверительных границ для дисперсии нормальной совокупности. Как известно, распределение выборочной дисперсии асимметрично, и мы можем сравнить точные результаты с даваемыми рассмотренной выше аппроксимацией.

Если определить

$$s^2 = \frac{1}{n}\sum(x - \bar{x})^2,$$

то известно, что для выборки объема n величина s^2/σ^2 имеет распределение типа III, иными словами, ns^2/σ^2 подчиняется χ^2-распределению с $n - 1$ степенями свободы.

Таким образом, для выборки объема 10 получаем (поскольку верхняя и нижняя 5-процентные точки для χ^2 равны соответственно 3,3251 и 16,9190)

$$\mathbf{P}\left\{3{,}3251 \leqslant \frac{10s^2}{\sigma^2} \leqslant 16{,}9190\right\} = 0{,}90.$$

Обращение этих неравенств дает

$$\mathsf{P}\{0{,}5911 s^2 \leqslant \sigma^2 \leqslant 3{,}001 s^2\} = 0{,}90. \tag{20.32}$$

Например, при $s^2 = 1$ доверительными границами служат 0,5911 и 3,001.

Беря $\theta = \sigma^2$, находим

$$\frac{\partial \log L}{\partial \theta} = \frac{n}{2\theta^2}\left\{\frac{1}{n}\sum (x-\mu)^2 - \theta\right\}, \tag{20.33}$$

откуда убеждаемся, что, как и требовалось,

$$\mathsf{M}\left(\frac{\partial \log L}{\partial \theta}\right) = 0. \tag{20.34}$$

Согласно примеру 17.10

$$I = \frac{n}{2\theta^2} \tag{20.35}$$

и, следовательно,

$$\frac{\partial I}{\partial \theta} = -\frac{n}{\theta^3}. \tag{20.36}$$

Дифференцируя (20.33) дважды и беря математическое ожидание, находим

$$\mathsf{M}\left(\frac{\partial^3 \log L}{\partial \theta^3}\right) = \frac{2n}{\theta^3}. \tag{20.37}$$

Таким образом, из (20.27), (20.36) и (20.37) имеем

$$\varkappa_3(J) = \frac{n}{\theta^3}. \tag{20.38}$$

Беря разложение (20.31) только до $\varkappa_3$, получаем

$$T(\theta) = \sqrt{\frac{2\theta^2}{n}}\left[\frac{n}{2\theta^2}\left\{\frac{1}{n}\sum (x-\mu)^2 - \theta\right\} - \right.$$
$$\left. - \frac{n}{6\theta^3}\frac{4\theta^4}{n}\left(\frac{n}{2\theta^2}\left\{\frac{n}{2\theta^2}\left(\frac{1}{n}\sum (x-\mu)^2 - \theta\right)^2\right\} - \frac{n}{2\theta^2}\right)\right] =$$
$$= \sqrt{\frac{2}{n}}\left[\frac{n}{2\theta}\left\{\frac{1}{n}\sum (x-\mu)^2 - \theta\right\} - \frac{n}{6\theta^2}\left\{\frac{1}{n}\sum (x-\mu)^2 - \theta\right\}^2 + \frac{1}{3}\right]. \tag{20.39}$$

Заменим $\sum (x-\mu)^2/n$ на величину $ns^2/(n-1)$, имеющую то же самое среднее значение. Тогда

$$T = \sqrt{\frac{2}{n}}\left[\frac{n}{2}\left(\frac{ns^2}{(n-1)\theta} - 1\right) - \frac{n}{6}\left(\frac{ns^2}{(n-1)\theta} - 1\right)^2 + \frac{1}{3}\right]. \tag{20.40}$$

Первый член дает нам доверительные границы для θ, основанные только на ψ. Два других являются поправочными членами более высокого порядка малости по n. Из (20.40) приближенно имеем

$$T = \sqrt{\frac{2}{n}}\left[\frac{n}{2}\left(\frac{ns^2}{(n-1)\theta} - 1\right) - \frac{n}{6}\cdot\frac{2}{n}T^2 + \frac{1}{3}\right],$$

откуда

$$\frac{ns^2}{(n-1)\theta} = 1 + T\sqrt{\frac{2}{n} + \frac{2}{3n}T^2} - \frac{2}{3n}. \qquad (20.41)$$

Например, если $n = 10$, $1 - \alpha = 0{,}90$, $s^2 = 1$ и $T = \pm 1{,}6449$ (5-процентные точки стандартного нормального распределения), то мы находим для s^2/θ границы 0,3403 и 1,6644, откуда для θ получаем границы 0,6008 и 2,939. Истинные значения, как мы знаем по (20.32), равны 0,5911 и 3,001. Для столь малого объема выборки, как $n = 10$, приближение кажется очень хорошим.

Кратчайшие множества доверительных интервалов*)

20.16 В примере 20.1 мы видели, что при некоторых условиях существует более одного множества доверительных интервалов. Нам необходимо понять, можно ли некоторое частное множество считать лучше других в каком-нибудь подходящем смысле. Аналогичная проблема возникала при оценивании, где в общем случае имеется много различных оценок для параметра, но иногда можно найти одну (как, например, с минимальной дисперсией), которая лучше всех остальных.

В примере 20.1 проблема имела несколько специализированную форму. Мы нашли, что для интервалов, основанных на значении $\bar{x}$, имеется бесконечное число множеств интервалов, каждое из которых соответствует способу выбора α_0 и α_1 (подчиненных условию $\alpha_0 + \alpha_1 = \alpha$). Среди них, очевидно, центральные интервалы являются кратчайшими, так как интервал данной длины содержит наибольшую долю нормального распределения, если его середина совпадает со средним значением. Разумно было бы сказать, что центральные интервалы являются наилучшими среди интервалов, определяемых с помощью $\bar{x}$.

Но из этого не следует, что они являются кратчайшими среди всех возможных интервалов или даже что кратчайшее множество существует. В общем случае для двух множеств интервалов c_1 и c_2 интервалы из c_1 могут быть короче, чем интервалы из c_2, для одних выборок и длиннее для других.

20.17 Поэтому мы будем рассматривать множества интервалов, которые являются кратчайшими в среднем. Это значит, что если

$$\delta = t_1 - t_0, \qquad (20.42)$$

*) Слово «множество» употребляется здесь с целью провести различие между «доверительным интервалом», отвечающим конкретной выборке, и правилом построения доверительных интервалов, дающим для всевозможных выборок «множество доверительных интервалов». (*Прим. ред.*)

то мы хотим минимизировать $\int \delta\, dF$, где интеграл берется по всем иксам и равен, следовательно,

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L\delta\, dx_1 \ldots dx_n. \tag{20.43}$$

Докажем сейчас теорему, принадлежащую Уилксу (1938b), которая очень сходна с результатом из **18.16** о том, что оценки максимального правдоподобия в пределе имеют минимальную дисперсию. Теорема эта говорит, что для больших выборок в определенном классе интервалов метод, приведенный в **20.10**, дает кратчайшие в среднем интервалы.

Пусть $h(x, \theta)$ — статистика с нулевым средним значением, и пусть сумма таких статистик подчиняется центральной предельной теореме. Это значит, что

$$\zeta = \frac{1}{\sqrt{n\,\mathrm{D}h}} \sum_{j=1}^{n} h(x_j, \theta) \tag{20.44}$$

имеет в пределе нормальное распределение со средним нуль и единичной дисперсией. Величина ψ из (20.8) является величиной типа ζ с $h = \partial \log f(x, \theta)/\partial\theta$. Покажем вначале, что абсолютная величина средней скорости изменения ψ относительно θ при каждом фиксированном θ больше, чем для любой ζ, исключая тривиальный случай $h = k\dfrac{\partial \log f}{\partial\theta}$. Полагая $g(x, \theta) = \dfrac{\partial \log f}{\partial\theta}$, имеем

$$\left.\begin{aligned} \frac{\partial\psi}{\partial\theta} &= \frac{1}{\sqrt{n\,\mathrm{D}g}}\left\{\sum \frac{\partial g}{\partial\theta} - \frac{1}{2\,\mathrm{D}g}\sum g\,\frac{\partial\,\mathrm{D}g}{\partial\theta}\right\}, \\ \frac{\partial\zeta}{\partial\theta} &= \frac{1}{\sqrt{n\,\mathrm{D}h}}\left\{\sum \frac{\partial h}{\partial\theta} - \frac{1}{2\,\mathrm{D}h}\sum h\,\frac{\partial\,\mathrm{D}h}{\partial\theta}\right\}. \end{aligned}\right\} \tag{20.45}$$

Следовательно,

$$\mathrm{M}\left(\frac{\partial\psi}{\partial\theta}\right) = \frac{1}{\sqrt{n\,\mathrm{D}g}}\left\{\sum \mathrm{M}\left(\frac{\partial g}{\partial\theta}\right) - \frac{1}{2\,\mathrm{D}g}\sum \mathrm{M}(g)\,\frac{\partial\,\mathrm{D}g}{\partial\theta}\right\}. \tag{20.46}$$

Далее, согласно (20.25) $\mathrm{M}(g) = 0$, а по (20.26)

$$\mathrm{M}\left(\frac{\partial g}{\partial\theta}\right) = \mathrm{M}\left(\frac{\partial^2 \log f}{\partial\theta^2}\right) = -\,\mathrm{M}\left(\frac{\partial \log f}{\partial\theta}\right)^2 = -\,\mathrm{D}g.$$

Поэтому (20.46) преобразуется следующим образом:

$$\mathrm{M}\left(\frac{\partial\psi}{\partial\theta}\right) = -\frac{n\,\mathrm{D}g}{\sqrt{n\,\mathrm{D}g}} = -\sqrt{n\,\mathrm{D}g}. \tag{20.47}$$

Аналогично, используя равенство $\mathrm{M}(h) = 0$, из (20.45) получаем

$$\mathrm{M}\left(\frac{\partial\zeta}{\partial\theta}\right) = \sqrt{\frac{n}{\mathrm{D}h}}\,\mathrm{M}\left(\frac{\partial h}{\partial\theta}\right). \tag{20.48}$$

Поскольку $0 = \mathbf{M}(h) = \int hf\,dx$, то, дифференцируя под знаком интеграла, находим

$$0 = \frac{\partial}{\partial\theta}\mathbf{M}(h) = \int \frac{\partial h}{\partial\theta} f\,dx + \int h\,\frac{\partial f}{\partial\theta}\,dx,$$

откуда

$$\mathbf{M}\left(\frac{\partial h}{\partial\theta}\right) = \int \frac{\partial h}{\partial\theta} f\,dx = -\int h\,\frac{\partial f}{\partial\theta}\,dx = -\operatorname{cov}(h,\ g). \quad (20.49)$$

Следовательно, из (20.47)—(20.49)

$$\left|\mathbf{M}\left(\frac{\partial\psi}{\partial\theta}\right)\right| - \left|\mathbf{M}\left(\frac{\partial\zeta}{\partial\theta}\right)\right| = \sqrt{\frac{n}{\mathbf{D}h}}\left\{\sqrt{\mathbf{D}h\,\mathbf{D}g} - |\operatorname{cov}(h,\ g)|\right\}. \quad (20.50)$$

Используя неравенство Коши — Буняковского, видим, что множитель в фигурных скобках в (20.50) положителен, если только h не равно константе, умноженной на g (тогда он обращается в нуль). Исключая этот случай, имеем

$$\left|\mathbf{M}\left(\frac{\partial\psi}{\partial\theta}\right)\right| > \left|\mathbf{M}\left(\frac{\partial\zeta}{\partial\theta}\right)\right|. \quad (20.51)$$

Это наш предварительный результат. Если теперь определить λ_α из соотношения

$$(2\pi)^{-1/2}\int_0^{\lambda_\alpha} \exp\left(-\frac{1}{2}x^2\right)dx = \frac{1}{2}(1-\alpha),$$

то доверительные границы для θ, скажем t_0 и t_1, полученные с помощью ψ, удовлетворяют уравнению $\sum g(x,\ \theta)/\sqrt{n\mathbf{D}h} = \pm\lambda_\alpha$, которое можно записать в виде $\psi(t) = \pm\lambda_\alpha$. Аналогично, границы, полученные с помощью ζ, скажем u_0 и u_1, удовлетворяют уравнению $\sum h(x,\ \theta)/\sqrt{n\mathbf{D}h} = \pm\lambda_\alpha$, которое перепишем в виде $\zeta(u) = \pm\lambda_\alpha$. Из разложения в ряд Тейлора в окрестности истинного значения параметра θ_0 получаем

$$\pm\lambda_\alpha = \psi(\theta_0) + (t-\theta_0)\left(\frac{\partial\psi}{\partial\theta}\right)_{\theta'} = \zeta(\theta_0) + (u-\theta_0)\left(\frac{\partial\zeta}{\partial\theta}\right)_{\theta''}, \quad (20.52)$$

где θ', θ'' — значения в окрестности θ_0, сходящиеся по вероятности к θ_0 при возрастании n. Полагая $t = u = \theta_0$ в (20.52), находим, что $\psi(\theta_0) = \zeta(\theta_0)$.

Следовательно,

$$(t-\theta_0)\left(\frac{\partial\psi}{\partial\theta}\right)_{\theta'} = (u-\theta_0)\left(\frac{\partial\zeta}{\partial\theta}\right)_{\theta''}. \quad (20.53)$$

Производные в (20.53) будут сходиться по вероятности к их математическим ожиданиям. Таким образом, из (20.51) и (20.53) для больших n имеем

$$|t-\theta_0| < |u-\theta_0|,$$

так что доверительные границы t_0 и t_1 ближе в среднем, чем любые другие границы u_0 и u_1, получаемые с помощью статистик вида (20.44).

20.18 Результат пункта **20.17** иллюстрирует тесную связь между теорией доверительных интервалов и теорией точечного оценивания, изложенной в главе 17. В **17.15—17** было показано, что МГД для оценки параметра θ равна $1/\mathbf{M}\{(\partial \log L/\partial\theta)^2\}$ и может быть достигнута лишь оценкой, являющейся линейной функцией от $\partial \log L/\partial\theta$. Естественно ожидать, что интервальная оценка параметра θ, основанная на $\partial \log L/\partial\theta$, будет обладать соответствующим свойством, а именно будет в среднем кратчайшей. Мы только что видели, что для больших выборок это действительно так.

20.19 Нейман (1937b) предложил использовать термин «кратчайшие доверительные интервалы» для множеств интервалов, определенных совершенно иным путем. Поскольку понятно, что такие интервалы не обязательно являются кратчайшими в смысле минимальной длины, даже в среднем, мы будем называть их «наиболее селективными», чтобы избежать путаницы в терминологии.

Рассмотрим множество интервалов s_0 с элементами δ_0, удовлетворяющими условию

$$\mathbf{P}\{\delta_0 c\theta \mid \theta\} = 1 - \alpha, \tag{20.54}$$

где $\delta_0 c\theta$ обозначает, что интервал δ_0 «содержит» θ, т. е. что $t_0 \leqslant \theta \leqslant t_1$. Пусть s_1 — некоторое другое множество с элементами δ_1 такими, что

$$\mathbf{P}\{\delta_1 c\theta \mid \theta\} = 1 - \alpha. \tag{20.55}$$

Для нас приемлемо каждое из этих множеств, так как в обоих случаях интервал δ содержит θ с вероятностью $1 - \alpha$.

Если теперь для любого s_1 мы при всяком значении θ', отличном от истинного значения параметра, имеем

$$\mathbf{P}\{\delta_0 c\theta' \mid \theta\} \leqslant \mathbf{P}\{\delta_1 c\theta' \mid \theta\}, \tag{20.56}$$

то множество интервалов s_0 будем называть *наиболее селективным*.

20.20 Идеи, лежащие в основе этого определения, будут разъяснены в главах 22 и 23, где рассматривается теория проверки гипотез. Забегая вперед, укажем, что смысл наиболее селективных интервалов состоит в том, что они покрывают истинное значение параметра с фиксированной вероятностью $1 - \alpha$, другие же значения параметра — с минимально возможной вероятностью. Относительно s_0 и s_1 можно сказать, что утверждение $\delta c \theta$ верно для них в $(1 - \alpha)$-й части случаев. Наиболее селективное множество s_0 выделяется тем, что оно покрывает ложные

значения параметра менее часто, чем их покрывают другие множества.

Различие между этим подходом и подходом, ведущим к кратчайшим интервалам, состоит в том, что последний имеет дело только с физической длиной доверительных интервалов, тогда как при первом минимизируют частоту, с которой покрываются альтернативные значения параметра θ. Один нацелен на локализацию истинного значения параметра θ с наибольшей гарантией от ошибки, другой принимает в расчет желание исключить из интервала как можно больше ложных значений θ, так чтобы ошибка принятия неверного значения была минимальной. Оказывается, что «селективный» подход математически проще, поэтому ему было уделено гораздо больше внимания. Относительно связи между двумя подходами см. ниже упражнения 20.17, 20.18. Мадански (1962) приводит пример, который выясняет различие между ними. Хартер (1964) обсуждает другие критерии для интервалов, предпочитая критерий, основанный на среднеквадратическом уклонении доверительных границ от истинного значения параметра.

Нейман показал, что наиболее селективные множества обычно не существуют (например, если распределение непрерывно), и предложил две альтернативные системы:

(а) наиболее селективные односторонние системы (по Нейману «кратчайшие односторонние» множества), подчиняющиеся (20.56) только для значений $\theta' - \theta$, которые всегда положительны или всегда отрицательны;

(б) селективные несмещенные системы (по Нейману «короткие несмещенные» множества), удовлетворяющие условию

$$\mathbf{P}\{\delta c\theta \mid \theta\} = 1 - \alpha \geqslant \mathbf{P}\{\delta c\theta \mid \theta'\}. \qquad (20.57)$$

Эти определения также представляют собой перевод некоторых идей из теории проверки гипотез на язык доверительных интервалов, и можно отложить их рассмотрение до главы 23. Поэтому в настоящей главе мы не будем проводить систематического изучения «оптимальных» доверительных интервалов.

20.21 Таблицы и графики для доверительных интервалов

(1) *Биномиальное распределение.* Клоппер и Пирсон (1934) опубликовали графики центральных доверительных интервалов для параметра распределения при $\alpha = 0{,}01$ и $0{,}05$; каждый из них дает контуры для $n = 8(2)12(4)24$, 30, 40, 60, 100, 200, 400 и 1000. Эти графики воспроизведены в *Biometrika Tables.* Для вдвое меньших значений α можно получить нижние или верхние односторонние интервалы. Можно также использовать таблицы неполной бета-функции (см. **5.7** и *Biometrika Tables*).

Пэчерс (1960) приводит центральные границы для $\alpha = 0{,}01$; 0,02; 0,05; 0,10 и $n = 55(5)100$ и ссылки на другие таблицы (в том числе на таблицы Кларка (1953)) для тех же значений α и $n = 10(1)50$ *).

Стерн (1954) предложил другой метод построения доверительных границ для вероятности успеха в случае биномиального распределения. Вместо того, чтобы быть центральной, доверительная полоса содержит значения p с наибольшими вероятностями их появления. Так как распределение p в общем случае асимметрично, то ясно, что таким путем мы интервал укорачиваем. Кроу (1956) показал, что эти интервалы образуют доверительную полосу с минимальной общей площадью, и табулировал несколько измененное множество интервалов для выборок объема до 30 и коэффициентов доверия 0,90; 0,95 и 0,99. См. также **20.23** ниже.

(2) *Распределение Пуассона.* (а) В *Biometrika Tables* приведены центральные доверительные интервалы для параметра распределения, построенные с помощью работы Гарвуда (1936) для наблюденных значений $x = 0(1)30(5)50$ и $\alpha = 0{,}002$; 0,01; 0,02; 0,05; 0,10. Как и в (1), это дает односторонние интервалы для $\alpha/2$ **). (б) Рикер (1937) приводит аналогичные таблицы для $x = 0(1)50$ и $\alpha = 0{,}01$; 0,05. (в) Пжиборовский и Виленский (1935) дают верхние доверительные границы только для $x = 0(1)50$, $\alpha = 0{,}001$; 0,005; 0,01; 0,02; 0,05; 0,10. (г) Кроу и Гарднер (1959) табулировали модифицированные интервалы Стерна — Кроу биномиального типа для $x = 0(1)300$ и $\alpha = 0{,}001$; 0,01; 0,05; 0,10; 0,20. См. также **20.23**.

(3) *Дисперсия нормального распределения.* Тейт и Клетт (1959) приводят наиболее селективные несмещенные доверительные интервалы и физически кратчайшие интервалы, основанные на достаточной статистике, для $\alpha = 0{,}001$; 0,005; 0,01; 0,05; 0,10 и $n = 3(1)30$. Первые из вышеупомянутых интервалов приводит также Пэчерс (1961) для $\alpha = 0{,}01$; 0,05; 0,10 и $n - 1 = 1(1)20$, 24, 30, 40, 60, 120 и Линдли и др. (1960) для $\alpha = 0{,}001$; 0,01; 0,05 и $n - 1 = 1(1)100$.

(4) *Отношение нормальных дисперсий.* Рамачандран (1958) приводит наиболее селективные несмещенные интервалы для $\alpha = 0{,}05$ и $n_1 - 1$, $n_2 - 1 = 2(1)4(2)12(4)24$, 30, 40, 60.

(5) *Коэффициент корреляции.* Ф. Дэвид (1938) приводит четыре графика центральных доверительных интервалов для

*) Центральные доверительные интервалы для параметра биномиального распределения даны в *Таблицах математической статистики* для $\alpha = 0{,}01$; 0,05 и 0,1 и $x, n - x = 0(1)20(2)30(5)50$, 60, 80, 100, 200, 500, где x — наблюденное число «успехов», и в *Сборнике статистических таблиц* для $\alpha = 0{,}01$; 0,02; 0,05; 0,1; 0,2; 0,5 и $x = 0(1)9$, $n - x = 1(1)12, 15(5)25$, 50, 100, 500. (*Прим. ред.*)

**) Для параметра распределения Пуассона в *Таблицах математической статистики* даны расширенные по сравнению с Biometrika Tables таблицы для $x = 0(1)50$ и поправки к приближенным формулам для вычисления доверительных границ при $x > 50$. (*Прим. ред.*)

коэффициента корреляции ρ двумерной нормальной совокупности для $\alpha = 0{,}01$; 0,02; 0,05; 0,10; на каждом из графиков даны контуры для $n = 3(1)8$, 10, 12, 15, 20, 25, 50, 100, 200 и 400. В *Biometrika Tables* воспроизведены указанные графики для $\alpha = 0{,}01$ и $\alpha = 0{,}05$ *). Односторонние интервалы могут быть получены, как и в (1).

Дискретность

20.22 При обсуждении биномиального распределения в примере 20.2 мы обратили внимание на тот факт, что число успехов (скажем, c) необходимо целое и доля успехов p $(=c/n)$, следовательно, дискретна. Полученная там доверительная полоса не была точной и давала доверительные утверждения вида $\mathsf{P} \geqslant 1 - \alpha$ вместо $\mathsf{P} = 1 - \alpha$. Прибегнув к специальному приему, можно всегда сделать утверждение типа $\mathsf{P} = 1 - \alpha$, даже в дискретном случае. Этот метод предложил Стивенс (1950).

Действительно, после того, как мы извлекли случайную выборку и получили c успехов, возьмем откуда-нибудь случайное число x, имеющее прямоугольное распределение $dF = dx$, $0 \leqslant x \leqslant 1$, например, выбирая случайным образом четыре цифры из обычных таблиц и принимая их за десятичные знаки. Тогда величина

$$y = c + x \tag{20.58}$$

может принимать любое значение в интервале от 0 до $n+1$ (предполагая, разумеется, что четырех десятичных знаков достаточно для того, чтобы задать значение непрерывной переменной). Если $y_0 = c_0 + x_0$, то, обозначая через ω оцениваемую вероятность, имеем

$$\mathsf{P}(y \geqslant y_0) = \mathsf{P}(c > c_0) + \mathsf{P}(c = c_0)\,\mathsf{P}(x \geqslant x_0) =$$

$$= \sum_{j=c_0+1}^{n} \binom{n}{j} \omega^j (1-\omega)^{n-j} + \binom{n}{c_0} \omega^{c_0} (1-\omega)^{n-c_0} (1 - x_0) =$$

$$= x_0 \sum_{j=c_0+1}^{n} \binom{n}{j} \omega^j (1-\omega)^{n-j} + (1 - x_0) \sum_{j=c_0}^{n} \binom{n}{j} \omega^j (1-\omega)^{n-j}. \tag{20.59}$$

Это определяет непрерывное распределение для y. Оно очевидным образом непрерывно, когда x изменяется от $0+$ до $1-$, так как c_0 тогда постоянно. В точках, где $x_0 = 0$, вероятность также стремится к одному и тому же значению как справа, так и слева. Можно теперь использовать это распределение, чтобы получать доверительные границы для ω, и доверительные утверждения,

*) Эти графики воспроизведены в *Таблицах математической статистики*. (*Прим. ред.*)

основанные на них, будут точными утверждениями типа $\mathsf{P} = 1 - \alpha$.

Доверительные интервалы будут иметь вид, показанный на рис. 20.6. Верхняя граница теперь сдвинулась вправо благодаря тому, что разрывы соединились последовательностью дуг. Нижняя граница также содержит последовательность дуг, но она не сдвинулась вправо; поэтому на рисунке показана (пунктиром) только приближенная верхняя граница из рис. 20.2. При нашем масштабе нижняя приближенная граница почти совпала бы с нижней последовательностью дуг. Общий результат состоит в том, что таким приемом доверительный интервал укорочен.

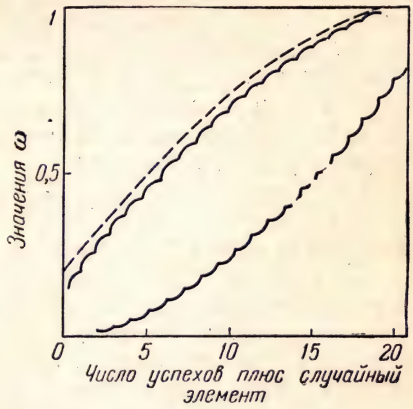


Рис. 20.6. Рандомизованные доверительные интервалы для биномиального параметра.

20.23 На первый взгляд удивительно, что интервалы, полученные таким путем, лежат внутри приближенных ступенчатых интервалов из рис. 20.2 и, следовательно, дают не менее точные границы. Ведь беря дополнительную случайную величину x, мы вносили дополнительную неопределенность в ситуацию. Однако небольшое размышление покажет, что то, что мы сделали, не бессмысленно. Мы устранили одну неопределенность, связанную с неравенством $\mathsf{P} \geqslant 1 - \alpha$, введением другой так, чтобы делать утверждения вида $\mathsf{P} = 1 - \alpha$, и потери от второй неопределенности более чем возмещаются устранением первой.

Центральные интервалы для параметра биномиального распределения можно легко получить из соотношения (20.59), но они не будут несмещенными наиболее селективными рандомизованными интервалами. Последние табулированы Блитом и Хатчинсоном (1960) для $\alpha = 0{,}01$; 0,05 и $n = 2(1)24(2)50$. Те же авторы (1961) приводят интервалы с такими же свойствами для параметра пуассоновского распределения при наблюдениях x, не превосходящих 250, и $\alpha = 0{,}01$; 0,05.

Обобщение на случай нескольких параметров

20.24 Теперь мы перейдем к обобщению предшествующей теории на случай распределения, зависящего от нескольких параметров. В дальнейшем, чтобы упростить изложение, мы будем подробно рассматривать лишь случай одномерной варианты,

хотя теория совершенно общая. Мы начнем с обозначений и с введения геометрической терминологии, которую можно рассматривать как обобщение понятий из рис. 20.1 и 20.2.

Предположим, что у нас есть плотность известного вида, зависящая от l неизвестных параметров $\theta_1, \ldots, \theta_l$. Обозначим ее $f(x, \theta_1, \ldots, \theta_l)$. Задача может состоять в оценке либо только θ_1, либо одновременно нескольких параметров θ_i. Вначале рассмотрим оценку лишь одного параметра. Для того чтобы определить доверительные границы, требуется найти две функции u_0 и u_1, зависящие от выборочных значений и не зависящие от параметров θ_i, такие, что

$$\mathbf{P}\{u_0 \leqslant \theta_1 \leqslant u_1\} = 1 - \alpha, \tag{20.60}$$

где $1 - \alpha$ есть заранее выбранный коэффициент доверия.

С выборкой $x_1, \ldots, x_n$ объема n можно связать точку в n-мерном евклидовом пространстве, а функция плотности будет определять плотность распределения для каждой такой точки. Величины u_0 и u_1, являясь функциями от иксов, определены в этом пространстве и для любого заданного коэффициента доверия $1 - \alpha$ лежат на двух гиперповерхностях (естественное обобщение понятия доверительных линий из рис. 20.1). Между ними заключена доверительная зона.

В общем случае мы должны рассматривать область априори возможных значений θ. Таким образом, к n-мерному пространству присоединяется l-мерное пространство значений θ. Общая область изменения имеет теперь размерность $l + n$. Но если мы оцениваем только параметр θ_1, то совокупное пространство редуцируется к $(n + 1)$-мерному пространству, остальные $l - 1$ параметров отсутствуют.

Будем обозначать через W выборочное пространство, а через E — точку с координатами $x_1, \ldots, x_n$. Будем писать $u_0(E)$, $u_1(E)$, чтобы подчеркнуть, что доверительные функции зависят от E. Интервал $u_1(E) - u_0(E)$ обозначим $\delta(E)$ или δ. Как и выше, $\delta c \theta_1$ означает, что $u_0 \leqslant \theta_1 \leqslant u_1$. Доверительную зону обозначим A и будем писать $E \in \delta$ или $E \in A$, чтобы показать, что выборочная точка лежит в интервале δ или в области A.

20.25 На рис. 20.7 показаны две оси x_1 и x_2 и третья ось, соответствующая переменной θ_1. Выборочное пространство W, таким образом, двумерно. Для любого заданного θ_1, скажем для θ_1', пространство W есть гиперплоскость (или ее часть), как показано на рисунке.

Возьмем любую заданную пару значений (x_1, x_2) и проведем через эту точку прямую, параллельную оси θ_1 (как PQ на рисунке) и пересекающую гиперплоскость в точке R. Соответствующие значения функций u_0 и u_1 будут давать две границы для θ_1; им отвечают две точки на этой прямой, например U и V.

Рассмотрим теперь прямые PQ, когда x_1 и x_2 изменяются. В некоторых случаях точки U, V находятся по разные стороны от R и параметр θ_1 лежит внутри интервала UV. В других случаях (как, например, для интервала $U'V'$, показанного на рисунке) получается обратное. Совокупность точек первой категории определяет область A, заштрихованную на рисунке. Для любой точки из A утверждение о том, что $\delta c\theta'$, является справедливым; для точек же, лежащих вне A, оно будет ошибочным.

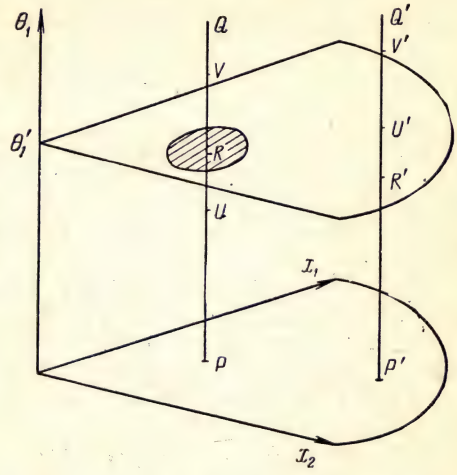


Рис. 20.7. Доверительные интервалы для $n = 2$ (см. текст).

20.26 Очевидно, если выборочная точка E попадает в область A, то соответствующее значение параметра θ_1 лежит в доверительном интервале, и наоборот. Из этого следует, что вероятность, с которой любое *фиксированное значение* θ_1 будет накрыто доверительным интервалом, равна вероятности того, что точка E лежит в $A(\theta_1)$, т. е.

$$\mathbf{P}\{\delta c\theta_1' \mid \theta_1, \ldots, \theta_l\} = \mathbf{P}\{u_0 \leqslant \theta_1' \leqslant u_1 \mid \theta_1, \ldots, \theta_l\} = \\ = \mathbf{P}\{E \in A(\theta_1') \mid \theta_1, \ldots, \theta_l\}. \quad (20.61)$$

Отсюда следует, что если доверительные функции определены так, что

$$\mathbf{P}\{u_0 \leqslant \theta_1 \leqslant u_1\} = 1 - \alpha,$$

то для любого θ_1

$$\mathbf{P}\{E \in A(\theta_1) \mid \theta_1, \ldots, \theta_l\} = 1 - \alpha. \quad (20.62)$$

Это показывает также, что ни при каком θ_1 область A не может быть пустой, так как вероятность в (20.62) была бы тогда равна нулю.

20.27 Если функции u_0 и u_1 однозначны и определены для всех E, то любая выборочная точка попадает по крайней мере в одну область $A(\theta_1')$. Действительно, если на прямой PQ, соответствующей данному E, возьмем точку R между U и V, то этим мы определим значение θ_1, например θ_1', такое, что $E \in A(\theta_1')$.

Более существенно то, что если выборочная точка попадает в области $A(\theta_1')$ и $A(\theta_1'')$, соответствующие двум значениям параметра θ_1, скажем θ_1' и θ_1'', то она попадает и в область $A(\theta_1''')$, где θ_1''' — любое значение между θ_1' и θ_1''. В самом деле, имеем $u_0 \leqslant \theta_1' \leqslant u_1$, $u_0 \leqslant \theta_1'' \leqslant u_1$, и, следовательно,

$$u_0 \leqslant \theta_1' \leqslant \theta_1''' \leqslant \theta_1'' \leqslant u_1,$$

если θ_1'' — наибольшее из θ_1' и θ_1''.

Далее, если выборочная точка попадает в каждую из областей $A(\theta_1)$, отвечающих значениям параметра θ, для которых $\theta_1' < \theta_1 < \theta_1''$, то она должна также попадать в области $A(\theta_1')$ и $A(\theta_1'')$.

20.28 Условия, указанные в двух предыдущих пунктах, являются необходимыми. Докажем сейчас, что они и достаточны, а именно: пусть для любого значения θ_1 существует определенная в выборочном пространстве W область A такая, что

(1) $\mathbf{P}\{E \in A(\theta_1) \mid \theta_1\} = 1 - \alpha$, каковы бы ни были значения параметров $\theta_1, \ldots, \theta_l$;

(2) для любого E существует по меньшей мере одно θ_1, например θ_1', такое, что $E \in A(\theta_1')$;

(3) если $E \in A(\theta_1')$ и $E \in A(\theta_1'')$, то $E \in A(\theta_1''')$ для любого θ_1''', лежащего между θ_1' и θ_1'';

(4) если $E \in A(\theta_1)$ для любого θ_1, удовлетворяющего условию $\theta_1' < \theta_1 < \theta_1''$, то $E \in A(\theta_1')$ и $E \in A(\theta_1'')$.

Тогда доверительные границы u_0 и u_1 для параметра θ определяются как нижняя и верхняя грани значений θ_1, для которых фиксированная выборочная точка попадает внутрь $A(\theta_1)$. Они определены и однозначны для всех E, $u_0 \leqslant u_1$ и $\mathbf{P}\{u_0 \leqslant \theta_1 \leqslant u_1 \mid \theta_1\} = 1 - \alpha$ при любом θ_1.

Нижняя и верхняя грани существуют вследствие условия (2), и нижняя грань не превосходит верхней. Нам остается только показать, что $\mathbf{P}\{u_0 \leqslant \theta_1 \leqslant u_1 \mid \theta_1\} = 1 - \alpha$, а для этого, вследствие условия (1), достаточно показать, что

$$\mathbf{P}\{u_0 \leqslant \theta_1 \leqslant u_1 \mid \theta_1\} = \mathbf{P}\{E \in A(\theta_1) \mid \theta_1\}. \qquad (20.63)$$

Мы уже знаем, что если $E \in A(\theta_1)$, то $u_0 \leqslant \theta_1 \leqslant u_1$, и наш результат будет установлен, если мы докажем обратное утверждение.

Пусть то, что $E \in A(\theta_1)$, когда $u_0 \leqslant \theta_1 \leqslant u_1$, — неверно. Пусть E — точка вне $A(\theta_1)$, для которой $u_0 \leqslant \theta_1 \leqslant u_1$. Тогда либо $u_0 = \theta_1$, либо $u_1 = \theta_1$, либо выполнено то и другое, так как в противном случае, поскольку u_0 и u_1 — границы тех значений θ_1, для которых E лежит в $A(\theta_1)$, существовали бы значения

θ_1' и θ_1'' такие, что $E \in A(\theta_1')$, $E \in A(\theta_1'')$ и

$$u_0 \leqslant \theta_1' \leqslant \theta_1 \leqslant \theta_1'' \leqslant u_1,$$

так что согласно условию (3) $E \in A(\theta_1)$, а это противоречит предположению.

Итак, либо $u_0 = \theta_1$, либо $u_1 = \theta_1$, либо выполнены оба равенства. В последнем случае точка E должна попасть в область $A(\theta_1)$, поскольку u_0 и u_1 — нижняя и верхняя границы значений θ, для которых $E \in A(\theta_1)$. Наконец, если $u_0 = \theta_1 < u_1$ (аналогично, если $u_0 < \theta_1 = u_1$), то мы видим, что для любого $u_0 < \theta_1 < u_1$ точка E должна по условию (3) попасть в $A(\theta_1)$ и, следовательно, по условию (4) E должна попасть в $A(\theta_1')$ и $A(\theta_1'')$, где $\theta_1' = u_0$ и $\theta_1'' = u_1$. Таким образом, она попадает в $A(\theta_1)$.

Выбор статистики

20.29 Приведенная выше теорема дает формальное решение задачи нахождения доверительных интервалов для одномерного параметра в общем случае. Однако она не содержит метода, с помощью которого можно находить интервал в конкретных примерах. На практике применяются четыре подхода: (1) использовать одномерную достаточную статистику, если она существует; (2) принять процесс, называемый «стьюдентизацией» (см. **20.31**); (3) «угадать» множество интервалов, исходя из общих знаний и опыта, и проверить, что оно удовлетворяет или не удовлетворяет требуемым условиям; (4) использовать аппроксимацию, основанную на обобщении метода из **20.15**.

20.30 Рассмотрим, как в общем случае используется одномерная достаточная статистика. Если t_1 — достаточная статистика для параметра θ_1, то

$$L = g(t_1 \mid \theta_1) L_2(x_1, \ldots, x_n, \theta_2, \ldots, \theta_l). \qquad (20.64)$$

Уравнения $t_1 = \text{const}$ определяют семейство гиперповерхностей в выборочном пространстве W. Если мы считаем, что этими гиперповерхностями задаются области в W, то, например, неравенство $t_1 \leqslant k$ задает некоторую фиксированную область K. Вероятность того, что E попадает в K, тогда зависит только от t_1 и θ_1. Соответствующим выбором k мы можем определить K так, что

$$\mathsf{P}\{E \in K \mid \theta_1\} = 1 - \alpha,$$

и, следовательно, получить области, основанные на значениях статистики t_1. Это можно сделать бесконечным числом способов, в зависимости от выбора значений α_0 и α_1. В **23.3** при

обсуждении данной задачи в терминах проверки гипотез мы увидим, что наиболее селективные интервалы (эквивалентные наиболее мощному критерию для $\theta_1 = \theta_1^0$) всегда могут быть построены с помощью достаточных статистик.

Стьюдентизация

20.31 В примере 20.1 мы рассматривали упрощенную задачу оценки среднего по выборке из нормальной совокупности с известной дисперсией. Предположим теперь, что нам требуется определить доверительные границы для среднего μ по выборке из совокупности

$$dF = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} dx,$$

где параметр σ неизвестен.

Рассмотрим величину $z = (\bar{x} - \mu)/s$, где s^2 — выборочная дисперсия. Как известно, она имеет распределение Стьюдента

$$dF = \frac{k\,dz}{(1+z^2)^{n/2}} \tag{20.65}$$

(см. пример 11.8). При заданном α мы можем теперь найти z_0 и z_1 такие, что

$$\int_{-\infty}^{-z_1} dF = \int_{z_1}^{\infty} dF = \frac{\alpha}{2},$$

откуда получаем соотношение $\mathbf{P}(-z_1 \leqslant z \leqslant z_0) = 1 - \alpha$, которое эквивалентно соотношению

$$\mathbf{P}(\bar{x} - sz_0 \leqslant \mu \leqslant \bar{x} + sz_1) = 1 - \alpha. \tag{20.66}$$

Следовательно, мы можем сказать, что μ лежит в интервале от $\bar{x} - sz_0$ до $\bar{x} + sz_1$ с коэффициентом доверия $1 - \alpha$, при этом интервал не зависит ни от μ, ни от σ. На самом деле, вследствие симметрии распределения Стьюдента, $z_0 = z_1$, но это побочное обстоятельство, несущественное для рассуждений.

Заметим, что выражение (20.66) (как и (20.4)) линейно относительно статистики $\bar{x}$; доверительными линиями в этом случае, как на рис. 20.1, будут параллельные прямые. Различие состоит в том, что при известном σ расстояние по вертикали между доверительными линиями является фиксированной функцией от σ, при неизвестном же параметре это расстояние есть случайная величина, поскольку оно зависит от s. Таким образом, здесь мы не можем фиксировать длину доверительного интервала заранее, до получения наблюдений.

20.32 Доверительные интервалы в данном случае удалось построить благодаря тому, что мы нашли статистику z, зависящую только от оцениваемого параметра, распределение ко-

торой не содержит σ. Таким путем часто можно устранить масштабный параметр, однако получаемое распределение не всегда легко вычислимо. Если, например, мы имеем статистику t степени p относительно переменных, то величина t/s^p имеет нулевую степень и ее распределение не зависит от масштабного параметра. Когда статистика таким путем преобразована к виду, не зависящему от масштабного параметра, говорят о «стьюдентизации», поскольку «Стьюдент» (Госсет) первым понял значение этого процесса.

20.33 Интересно рассмотреть связь между стьюдентизованной статистикой для среднего и доверительными зонами, основанными на достаточных статистиках, в случае нормального распределения. Совместное распределение среднего и дисперсии в выборке из нормальной совокупности есть (пример 11.7)

$$dF = \left(\frac{n}{2\pi\sigma^2}\right)^{1/2} \exp\left\{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2\right\} d\bar{x}\, \frac{k}{\sigma^{n-1/2}}\, s^{n-3} \exp\left\{-\frac{ns^2}{2\sigma^2}\right\} ds^2, \tag{20.67}$$

и $\bar{x}$, s совместно достаточны (пример 17.17). В выборочном пространстве W области, где $\bar{x}$ постоянно, образуют гиперплоскости, а области с постоянной s — гиперсферы. Если мы фиксируем $\bar{x}$ и s, то выборочная точка E лежит на $(n-2)$-мерной гиперсфере (пример 11.7). Выберем на этой гиперсфере область, содержащую $(1-\alpha)$-часть от площади гиперсферы. Тогда доверительная зона A будет получена объединением таких областей при всех $\bar{x}$ и s.

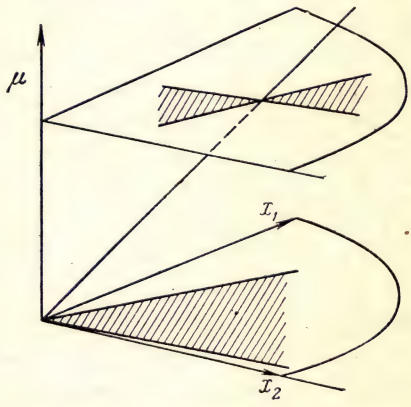


Рис. 20.8. Доверительные интервалы, основанные на t-статистике Стьюдента для $n=2$ (см. текст).

Одной из таких областей является «ломтик» выборочного пространства, полученный поворотом гиперплоскости, проходящей через начало координат и точку $(1, 1, \ldots, 1)$, на угол $\pi(1-\alpha)$ (а не на $2\pi(1-\alpha)$, потому что полуоборот этой плоскости покрывает все пространство).

На рис. 20.8 проиллюстрирован случай $n=2$.

Для любого заданного μ' ось вращения пересекает гиперплоскость $\mu=\mu'$ в точке $x_1=x_2=\mu'$, а гиперконусы $(\bar{x}-\mu)/s=\text{const}$ в пространстве W представляют собой плоские области, заключенные между двумя прямыми (на рисунке

заштриховано). Множество областей A получается с помощью такого поворота плоскости вокруг прямой $x_1 = x_2 = \mu$, при котором на любой плоскости $\mu = \mu'$ образуется угол $\frac{1}{2}\pi(1-\alpha)$ с каждой стороны от прямой $x_1 - \mu' = x_2 - \mu'$.

Граничные плоскости задаются уравнениями

$$x_1 - \mu = (x_2 - \mu)\operatorname{tg}\left(\frac{1}{4}\pi - \frac{1}{2}\beta\right),$$

$$x_1 - \mu = (x_2 - \mu)\operatorname{tg}\left(\frac{1}{4}\pi + \frac{1}{2}\beta\right),$$

где $\beta = \pi\alpha$, или, после небольших преобразований,

$$\mu = \frac{1}{2}(x_1 + x_2) + \frac{1}{2}(x_1 - x_2)\operatorname{ctg}\frac{1}{2}\beta,$$

$$\mu = \frac{1}{2}(x_1 + x_2) - \frac{1}{2}(x_1 - x_2)\operatorname{ctg}\frac{1}{2}\beta.$$

Тогда для μ мы получаем интервал

$$\frac{1}{2}(x_1 + x_2) - \frac{1}{2}|x_1 - x_2|\operatorname{ctg}\frac{1}{2}\beta \leqslant \mu \leqslant$$

$$\leqslant \frac{1}{2}(x_1 + x_2) + \frac{1}{2}|x_1 - x_2|\operatorname{ctg}\frac{1}{2}\beta.$$

Но это и есть границы, получаемые из распределения Стьюдента для $n = 2$. Действительно, выборочное стандартное отклонение равно $\frac{1}{2}|x_1 - x_2|$ и

$$\frac{1}{\pi}\int_{z_0}^{\infty}\frac{dz}{1+z^2} = \frac{1}{\pi}\left(\frac{\pi}{2} - \operatorname{arctg} z_0\right) = \frac{\alpha}{2} = \frac{\beta}{2\pi},$$

откуда

$$z_0 = \operatorname{tg}\left(\frac{\pi}{2} - \frac{\beta}{2}\right) = \operatorname{ctg}\frac{\beta}{2}.$$

20.34 Следующий пример иллюстрирует использование «стьюдентизации» при построении доверительных интервалов, а также результаты, к которым она может привести.

Пример 20.7. Доверительные интервалы для отношения средних двух нормальных величин

Пусть величины x, y имеют совместно нормальное распределение со средними ξ, η и неизвестными дисперсиями и ковариацией. Предположим, что ξ достаточно велико для того, чтобы область изменения x была существенно положительной.

Найдем доверительный интервал для отношения $\theta = \eta/\xi$, основанный на статистике $\bar{y}/\bar{x}$. Имеем

$$\mathsf{P}\left(\frac{y}{x} \leqslant \theta\right) = \mathsf{P}(y - \theta x \leqslant 0). \tag{20.68}$$

Величина $y - \theta x$ сама распределена нормально, и в выборке объема n среднее $\bar{y} - \theta\bar{x}$ распределено тоже нормально с дисперсией $(\mathsf{D}y - 2\theta \operatorname{cov}(x, y) + \theta^2 \mathsf{D}x)/n$. Следовательно (см. **16.10**), отношение

$$t = \frac{(\bar{y} - \theta\bar{x})\sqrt{n-1}}{\{\hat{\mathsf{D}}y - 2\theta\,\widehat{\operatorname{cov}}(x, y) + \theta^2 \hat{\mathsf{D}}x\}^{1/2}} \tag{20.69}$$

имеет t-распределение Стьюдента с $n-1$ степенями свободы, если для членов в знаменателе используются оценки вида $\sum (x - \bar{x})^2/n$.

Этот результат принадлежит Филлеру (1940). По таблицам, обычным путем, можно найти критические значения для величины t, и вопрос состоит лишь в том, можно ли, используя (20.69), указать соответствующие границы для θ. В данном случае не имеется одномерной достаточной статистики для параметра θ. Наше уравнение содержит пять достаточных статистик (два средних, выборочные дисперсии и ковариация), которые в определенной степени являются зависимыми. Можно ожидать, что некоторые трудности, с которыми мы сталкивались в **20.12—14**, возникнут и здесь. Так оно и есть.

Рассмотрим, как изменяются θ и t^2 при заданных значениях пяти статистик. Имеем

$$\frac{t^2}{n-1} \equiv \frac{\bar{y}^2\,\hat{\mathsf{D}}x - 2\bar{x}\bar{y}\,\widehat{\operatorname{cov}}(x, y) + \bar{x}^2\,\hat{\mathsf{D}}y}{\hat{\mathsf{D}}x\,\hat{\mathsf{D}}y - \{\widehat{\operatorname{cov}}(x, y)\}^2} -$$

$$- \frac{[\{\bar{y}\,\widehat{\operatorname{cov}}(x, y) - \bar{x}\,\hat{\mathsf{D}}y\} - \theta\{\bar{y}\,\hat{\mathsf{D}}x - \bar{x}\,\widehat{\operatorname{cov}}(x, y)\}]^2}{[\hat{\mathsf{D}}x\,\hat{\mathsf{D}}y - \{\widehat{\operatorname{cov}}(x, y)\}^2]\{\hat{\mathsf{D}}y - 2\theta\,\widehat{\operatorname{cov}}(x, y) + \theta^2\hat{\mathsf{D}}x\}}. \tag{20.70}$$

Уравнение (20.70) является уравнением третьего порядка относительно θ и t^2. Если откладывать θ как ординату, а t^2 как абсциссу, то получим график, аналогичный приведенному на рис. 20.9 (последний, как легко видеть, *не* изображает доверительных линий).

Максимальное значение t^2 (скажем, $t^2_{\max}$) получится из (20.70), когда

$$\theta = \frac{\bar{y}\,\widehat{\operatorname{cov}}(x, y) - \bar{x}\hat{\mathsf{D}}y}{\bar{y}\hat{\mathsf{D}}x - \bar{x}\,\widehat{\operatorname{cov}}(x, y)}. \tag{20.71}$$

Обозначим это значение через A. Минимальное значение достигается при $t^2 = 0$. Прямая $t^2/(n-1) = \bar{x}^2/\hat{\mathsf{D}}x$ является асимптотой.

Для $t^2 = 0$ или $t^2_{\max}$ оба значения θ совпадают. Когда $t^2 > t^2_{\max}$, они являются мнимыми. Если $0 < t^2 < A$, то они действительны и различны. При изменении t^2 от 0 до $t^2_{\max}$ больший корень θ монотонно возрастает (или убывает) от наблюденного значения $\bar{y}/\bar{x}$ до A. Меньший корень убывает (или возрастает) от $\bar{y}/\bar{x}$, уходит в бесконечность в асимптоте, вновь появляется с противоположным знаком с другой стороны асимптоты и монотонно приближается к A, снова сливаясь там с другим корнем. Границы для θ, соответствующие данному критическому значению t^2, указаны на рис. 20.9 (получено из работы Филлера, 1954). Для конкретных значений $\bar{x}$, $\bar{y}$, $\hat{\mathrm{D}}x$, $\widehat{\mathrm{cov}}(x, y)$ и $\hat{\mathrm{D}}y$ мы можем утверждать, что параметр θ лежит внутри действительного интервала при коэффициенте доверия, определяющем величину $t^2/(n-1)$, принадлежащую интервалу от 0 до $\bar{x}^2/\hat{\mathrm{D}}x$; что он лежит в интервале, бесконечном в отрицательном направлении, при $\bar{x}^2/\hat{\mathrm{D}}x \leqslant t^2/(n-1) < A$, и можем утверждать только, что он лежит где-то в интервале от $-\infty$ до $+\infty$, для $t^2/(n-1) > A$.

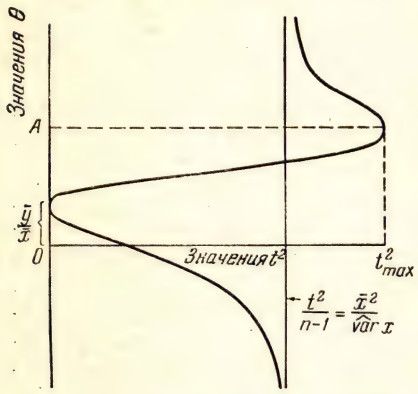


Рис. 20.9. Доверительные интервалы, основанные на (20.71) (см. текст).

Совместные доверительные интервалы для нескольких параметров

20.35 Довольно часто возникают задачи, где нужно оценить более чем один параметр, например среднее и дисперсию. Обобщение теории доверительных интервалов для одного параметра на случай двух или большего числа параметров наталкивается на существенные трудности. Пусть, скажем, даны два параметра θ_1 и θ_2 и две статистики t и u. Хотелось бы получить совместное интервальное утверждение типа

$$\mathrm{P}\{t_0 \leqslant \theta_1 \leqslant t_1 \quad \text{и} \quad u_0 \leqslant \theta_2 \leqslant u_1\} = 1 - \alpha. \tag{20.72}$$

К сожалению, это редко удается. Иногда можно сделать утверждение, определяющее *доверительную область* одновременно для двух параметров, как, например,

$$\mathrm{P}\{w_0 \leqslant \theta_1^2 + \theta_2^2 \leqslant w_1\} = 1 - \alpha. \tag{20.73}$$

Но это не вполне удовлетворительно, так как мы не знаем, так сказать, какая доля неопределенности этой области связана с каждым из параметров. Возможно, что без принятия какого-то нового подхода задача локализации параметров в отдельные интервалы неразрешима.

Проблема остается сложной даже при больших выборках. Можно определить интервалы типа

$$\mathbf{P}\{t_0(\theta_2) \leqslant \theta_1 \leqslant t_1(\theta_2)\} = 1 - \alpha$$

и подставить (при большой выборке) оценку для θ_2 в границы $t_0(\theta_2)$ и $t_1(\theta_2)$. Этот прием аналогичен хорошо известной процедуре в теории стандартных ошибок, где мы заменяем параметры, входящие в дисперсии ошибок, их оценками, полученными по выборкам.

20.36 Мы не будем пытаться развить здесь далее теорию совместных доверительных интервалов. Читатель, который интересуется теоретическим аспектом, может посмотреть работы С. Роя и Бозе (1953) и С. Роя (1954). Бартлетт (1953, 1955) рассмотрел обобщение метода из **20.15** на случай двух или более неизвестных параметров. Гальперин и Мантел (1963) и Гальперин (1964) рассмотрели интервалы для нелинейных функций от параметров (в частности, при больших выборках).

Теорема из **20.17**, касающаяся кратчайших интервалов, была обобщена Уилксом и Дейли (1939). При большой выборке в довольно общих предположениях наименьшая в среднем область для l параметров дается соотношением

$$\sum_{i=1}^{l}\sum_{j=1}^{l}\left\{ I_{ij}^{-1}\,\frac{\partial \log L}{\partial \theta_i}\,\frac{\partial \log L}{\partial \theta_j}\right\} \leqslant \chi_\alpha^2, \tag{20.74}$$

где I^{-1} — матрица, обратная к информационной матрице, общий элемент которой имеет вид

$$I_{ij} = \mathbf{M}\left(\frac{\partial \log L}{\partial \theta_i}\,\frac{\partial \log L}{\partial \theta_j}\right),$$

а χ_α^2 таково, что $\mathbf{P}(\chi^2 \leqslant \chi_\alpha^2) = 1 - \alpha$, где вероятность вычисляется с помощью χ^2-распределения с l степенями свободы. Это, очевидно, имеет отношение к результату из **17.39**, дающему минимальные достижимые дисперсии (а после простого обобщения и ковариации) для множества несмещенных оценок нескольких параметрических функций.

В третьем томе в разделе дисперсионного анализа нам встретится задача нахождения доверительных интервалов одновременно для нескольких средних.

Толерантные интервалы

20.37 В этой главе мы занимались вопросом нахождения доверительных интервалов для параметров, участвующих явно в задании распределения. Но техника доверительных интервалов может быть использована и для других задач. В следующих главах мы увидим, что можно находить интервалы для квантилей исходных распределений (см. упражнение 20.17), а также для самих функций распределения без каких-либо предположений о типе распределения, только в предположении непрерывности. Имеется и другой класс задач, часто встречающихся в практике, которые можно решить этими методами. Пусть на основе n независимых наблюдений из некоторого распределения мы хотим найти две границы L_1 и L_2, о которых можно утверждать, что доля распределения, содержащаяся между ними, не меньше чем γ. Понятно, что такие утверждения нужно делать в вероятностной форме, а именно, с заданной вероятностью β доля распределения, не меньшая γ, лежит между L_1 и L_2. Величины L_1 и L_2 называются *толерантными границами* для распределения; мы их будем называть (β, γ)-толерантными границами. Интервал (L_1, L_2) является *толерантным интервалом*. В главе 32 мы увидим, что толерантные границы можно строить без предположений (кроме непрерывности) о виде исходного распределения. В настоящей главе, однако, мы рассмотрим вывод толерантных границ для нормального распределения, принадлежащий Вальду и Волфовицу (1946).

20.38 Так как выборочные среднее и дисперсия образуют пару достаточных статистик для параметров нормального распределения (упражнение 17.17), то естественно на их основе искать толерантные границы для исходного распределения. По выборке объема n построим несмещенные оценки

$$\bar{x} = \sum x/n, \qquad s'^2 = \sum (x - \bar{x})^2/(n-1)$$

и положим

$$A(\bar{x}, s', \lambda) = \int_{\bar{x}-\lambda s'}^{\bar{x}+\lambda s'} f(t)\, dt, \tag{20.75}$$

где $f(t)$ — плотность нормального распределения. Требуется найти величину λ, для которой

$$\mathsf{P}\{A(\bar{x}, s', \lambda) > \gamma\} = \beta. \tag{20.76}$$

Тогда $L_1 = \bar{x} - \lambda s'$ и $L_2 = \bar{x} + \lambda s'$ образуют пару центральных (β, γ)-толерантных границ для исходного распределения. Так как нас интересует лишь доля распределения, покрываемая ин-

тервалом (L_1, L_2), то без потери общности можно считать, что истинное среднее равно 0, а дисперсия 1. Тогда

$$f(t) = (2\pi)^{-1/2} \exp\left(-\frac{1}{2}t^2\right). \tag{20.77}$$

20.39 Рассмотрим условную вероятность того, что $A(\bar{x}, s', \lambda)$ превзойдет γ, при заданном $\bar{x}$. Обозначим ее $\mathbf{P}\{A > \gamma \mid \bar{x}\}$. Функция A монотонно возрастает по s', и уравнение

$$A(\bar{x}, s', \lambda) = \gamma \tag{20.78}$$

имеет в точности один корень, который обозначим $s'(\bar{x}, \gamma, \lambda)$. Пусть

$$\lambda s'(\bar{x}, \gamma, \lambda) = r(\bar{x}, \gamma). \tag{20.79}$$

При заданных $\bar{x}$ и γ величину $r = r(\bar{x}, \gamma)$ можно сразу найти по таблицам нормального распределения, так как

$$\int_{\bar{x}-r}^{\bar{x}+r} f(t)\, dt = \gamma. \tag{20.80}$$

Из соотношения (20.80) видно, что r не зависит от λ. Кроме того, поскольку A монотонно возрастает относительно s', то неравенство $A > \gamma$ эквивалентно неравенству

$$s' > s'(\bar{x}, \gamma, \lambda) = r(\bar{x}, \gamma)/\lambda.$$

Таким образом, можно написать

$$\mathbf{P}\{A > \gamma \mid \bar{x}\} = \mathbf{P}\left\{s' > \frac{r}{\lambda} \,\middle|\, \bar{x}\right\}. \tag{20.81}$$

Так как $\bar{x}$ и s' независимы, то (20.81) переходит в соотношение

$$\mathbf{P}\{A > \gamma \mid \bar{x}\} = \mathbf{P}\{(n-1)s'^2 > (n-1)r^2/\lambda^2\}. \tag{20.82}$$

Поскольку величина $(n-1)s'^2 = \sum (x - \bar{x})^2$ имеет χ^2-распределение с $n-1$ степенями свободы, то окончательно имеем

$$\mathbf{P}\{A > \gamma \mid \bar{x}\} = \mathbf{P}\{\chi^2_{n-1} > (n-1)r^2/\lambda^2\}. \tag{20.83}$$

Используя таблицы χ^2-распределения, мы можем определить вероятность в соотношении (20.83).

20.40 Для того чтобы из (20.83) получить безусловную вероятность $\mathbf{P}(A > \gamma)$, мы должны это соотношение проинтегрировать по распределению $\bar{x}$, которое нормально со средним нуль и дисперсией $1/n$. Интегрирование является утомительной вычислительной операцией, но, к счастью, существует красивый приближенный метод. Разложим функцию $\mathbf{P}(A > \gamma \mid \bar{x})$ в ряд

Тейлора в окрестности точки $\bar{x} = \mu = 0$, и поскольку она является четной по $\bar{x}$ функцией, то нечетные степени в разложении будут отсутствовать *). Получим

$$\mathbf{P}(A > \gamma \mid \bar{x}) = \mathbf{P}(A > \gamma \mid 0) + \frac{\bar{x}^2}{2!}\mathbf{P}''(A > \gamma \mid 0) + \ldots \quad (20.84)$$

Беря математическое ожидание, имеем

$$\mathbf{P}(A > \gamma) = \mathbf{P}(A > \gamma \mid 0) + \frac{1}{2n}\mathbf{P}''(A > \gamma \mid 0) + \ldots \quad (20.85)$$

Но из (20.84) при $\bar{x} = 1/\sqrt{n}$ находим

$$\mathbf{P}\left(A > \gamma \,\middle|\, \frac{1}{\sqrt{n}}\right) = \mathbf{P}(A > \gamma \mid 0) + \frac{1}{2n}\mathbf{P}''(A > \gamma \mid 0) + O(n^{-2}). \quad (20.86)$$

Соотношения (20.85) и (20.86) дают

$$\mathbf{P}(A > \gamma) = \mathbf{P}\left(A > \gamma \,\middle|\, \frac{1}{\sqrt{n}}\right) + O(n^{-2}), \quad (20.87)$$

и мы можем использовать (20.87) для нахождения приближенного значения величины λ в (20.83). Вальд и Волфовиц (1946) показали (см. также Эллисон (1964)), что это приближение очень хорошее даже для таких малых значений n, как $n = 2$, если только β и $\gamma \geqslant 0{,}95$, каковыми они обычно бывают на практике. Боукер (1947) приводит таблицы значений λ (у него k) для β (у него γ) $= 0{,}75$; 0,90; 0,95; 0,99 и γ (у него P) $= 0{,}75$; 0,90; 0,99 и 0,999, для объемов выборок $n = 2(1)102(2)180(5)300(10)400(25)750(50)1000$ **).

Легко видеть, что приведенные рассуждения остаются справедливыми, если $\bar{x}$ заменить любой оценкой $\hat{\mu}$ для среднего, а s'^2 — любой независимой оценкой $\hat{\sigma}^2$ для дисперсии нормальной совокупности (см. Уоллис (1951)). Для случая, когда среднее оценено по n наблюдениям, а оценка дисперсии имеет ν степеней свободы, у Тагути (1958) приведены таблицы значений λ (у него k) при β (у него $1 - \alpha$) и γ (у него P) $= 0{,}90$; 0,95 и 0,99 и $n = 0{,}5(0{,}5)2(1)10(2)20(5)30(10)60(20)100$, 200, 500, 1000, ∞; $\nu = 1(1)20(2)30(5)100(100)1000$, ∞. Малые дробные значения n полезны для некоторых приложений, которые можно найти у Тагути. Вайсберг и Битти (1960) приводят таблицы для λ/r (у них u) при ν (у них f) $= 1(1)150(2)250(5)500(10)1000(1000)10000$, ∞ и β (у них γ) $= 0{,}90$; 0,95, 0,99 и для r

*) Данный факт верен вследствие симметрии интервала относительно $\bar{x}$. В отсутствие симметрии этого не могло быть.

**) Сокращенный вариант этих таблиц приведен в *Таблицах математической статистики*. (*Прим. ред.*)

при $n = 1(1)100(5)200(10)300(20)500(100)1000(1000)10000$, ∞ и γ (у них $P) = 0{,}5$; 0,75, 0,9; 0,95; 0,99; 0,999 *).

Фрэзер и Гаттмэн (1956) и Гаттмэн (1957) рассмотрели толерантные интервалы, которые покрывают *в среднем* заданную часть исходного нормального распределения.

УПРАЖНЕНИЯ

20.1 Мы видели (упражнение 17.1), что при выборке объема n из распределения

$$dF = \frac{x^{p-1}e^{-x/\theta}}{\Gamma(p)\,\theta^p}\,dx, \qquad 0 \leqslant x \leqslant \infty, \qquad p > 0,$$

когда p известно, достаточной статистикой для параметра θ является $\bar{x}/p$. Построить доверительный интервал для θ.

20.2 Показать, что в случае прямоугольного распределения

$$dF = dx/\theta, \qquad 0 \leqslant x \leqslant \theta,$$

при коэффициенте доверия $1 - \alpha$ доверительными границами для параметра θ служат t и t/ψ, где t — выборочный размах, а ψ задается уравнением

$$\psi^{n-1}\{n - (n-1)\,\psi\} = \alpha.$$

(Уилкс, 1938c.)

20.3 Показать, что в случае распределения из предыдущего упражнения доверительные границы для θ при выборке объема два, x_1 и x_2, суть

$$(x_1 + x_2)/[1 \pm \{1 - (1-\alpha)^{1/2}\}].$$

(Нейман, 1937b.)

20.4 В условиях упражнения 20.2 показать, что если L — большее наблюдение в выборке объема два, то доверительными границами для θ служат L, $L/\sqrt{\alpha}$, а если M — наибольшее наблюдение в выборке объема четыре, то границами служат

$$M, \quad M/\alpha^{1/4}.$$

(Нейман, 1937b.)

20.5 Используя асимптотическую многомерную нормальность оценок максимального правдоподобия (**18.26**) и χ^2-распределение показателя степени многомерного нормального распределения (**15.10**), показать, что соотношение (20.74) дает, при большой выборке, доверительную область для совокупности параметров. На основе этого получить доверительную область для среднего и дисперсии одномерного нормального распределения.

20.6 При построении доверительных границ для дисперсии нормальной совокупности с помощью распределения выборочной дисперсии (пример 20.6) начертить доверительные полосы для какого-нибудь значения коэффициента доверия и показать, что они всегда дают связный интервал, внутри которого заключена дисперсия σ^2.

---

*) Эта таблица воспроизведена в *Сборнике статистических таблиц*. (*Прим. ред.*)

20.7 Показать, как построить доверительные границы для отношения дисперсий σ_1^2/σ_2^2 двух нормальных совокупностей, если даны n_1 независимых наблюдений из первой совокупности и n_2 — из второй. (Использовать распределение отношения выборочных дисперсий (16.24).)

20.8 Используя метод из **20.10**, показать, что при большой выборке 95-процентные доверительные границы для ω в биномиальном распределении примера 20.2 имеют вид

$$\frac{1}{1+(1{,}96)^2/n}\left\{p+\frac{(1{,}96)^2}{2n}\pm 1{,}96\sqrt{\frac{p(1-p)}{n}+\frac{(1{,}96)^2}{4n^2}}\right\}.$$

20.9 Используя теорему Гири (упражнение 11.11), показать, что при большой выборке 95-процентные доверительные границы для отношения ω_2/ω_1 параметров двух биномиальных распределений, основанные на независимых выборках объемов n_2 и n_1 соответственно, имеют вид

$$\frac{p_2/p_1}{1+(1{,}96)^2/n_2}\left\{1+\frac{(1{,}96)^2}{2n_2p_2}+\right.$$
$$\left.+1{,}96\sqrt{\frac{1-p_1}{n_1p_1}+\frac{1-p_2}{n_2p_2}+\frac{(1{,}96)^2}{4}\left(\frac{1}{n_2^2p_2^2}+\frac{4(1-p_1)}{n_1n_2p_1}\right)}\right\}.$$

(Нётер, 1957.)

20.10 В условиях примера 20.6 показать, что доверительный интервал, основанный на

$$\mathrm{P}\left\{\frac{ns^2}{\chi_0^2}\leqslant\sigma^2\leqslant\frac{ns^2}{\chi_1^2}\right\}=1-\alpha$$

(где χ_0^2 и χ_1^2 — нижняя и верхняя $\frac{1}{2}\alpha$- точки χ^2-распределения с $n-1$ степенями свободы), не является физически кратчайшим интервалом для σ^2, построенным на основе χ^2-распределения величины ns^2/σ^2.

(см. Тейт и Клетт, 1959.)

20.11 Извлекаются независимые выборки объемов n_1 и n_2 соответственно из двух нормальных совокупностей со средними μ_1 и μ_2 и дисперсиями $\sigma_1^2=\sigma_2^2=\sigma^2$. Показать, что величина

$$t=\{(\bar{x}_1-\mu_1)-(\bar{x}_2-\mu_2)\}\Big/\left\{\frac{n_1s_1^2+n_2s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right\}^{1/2}$$

(где $\bar{x}_1$, $\bar{x}_2$ и s_1^2, s_2^2 — выборочные средние и дисперсии) имеет распределение Стьюдента с n_1+n_2-2 ст. св., и построить доверительные границы для $\mu_1-\mu_2$.

20.12 Показать, что если в упражнении 20.11 $\sigma_1^2\neq\sigma_2^2$, то статистикой, имеющей распределение Стьюдента, является не t, а

$$t'=\frac{(\bar{x}_1-\mu_1)-(\bar{x}_2-\mu_2)}{(\sigma_1^2/n_1+\sigma_2^2/n_2)^{1/2}}\Big/\left\{\left(\frac{n_1s_1^2}{\sigma_1^2}+\frac{n_2s_2^2}{\sigma_2^2}\right)\Big/(n_1+n_2-2)\right\}^{1/2}.$$

20.13 Показать, что если $f(x|\theta)=g(x)/h(\theta)$, $a(\theta)\leqslant x\leqslant b(\theta)$, и $b(\theta)$ — монотонно убывающая функция от $a(\theta)$, то экстремальные наблюдения $x_{(1)}$ и $x_{(n)}$ образуют пару совместно достаточных статистик для θ (см. **17.40—41**).

Используя их совместное распределение, показать, что одномерная достаточная статистика для параметра θ, $\hat{\theta} = \min\left\{a^{-1}(x_{(1)}),\ b^{-1}(x_{(n)})\right\}$, имеет распределение

$$dF = \frac{n\{h(\hat{\theta})\}^{n-1}}{\{h(\theta)\}^{n}}\{-h'(\hat{\theta})\}\,d\hat{\theta}, \qquad \theta \leqslant \hat{\theta} \leqslant \theta^*,$$

где θ^* определяется из уравнения $a(\theta^*) = b(\theta^*)$.

20.14 В условиях упражнения 20.13 показать, что величина $\psi = h(\hat{\theta})/h(\theta)$ имеет распределение $dF = n\psi^{n-1}\,d\psi$, $0 \leqslant \psi \leqslant 1$. Доказать, что

$$\mathbf{P}\{\alpha^{1/n} \leqslant \psi \leqslant 1\} = 1 - \alpha,$$

и, пользуясь этим, построить доверительный интервал для θ. Доказать, что он короче любого другого интервала, основанного на распределении величины ψ.

(Хузурбазар, 1955.)

20.15 Используя результат упражнения 20.14, показать, что доверительный интервал для параметра θ в случае распределения

$$dF = \frac{dx}{\theta}, \qquad 0 \leqslant x \leqslant \theta,$$

получается из соотношения

$$\mathbf{P}\left\{x_{(n)} \leqslant \theta \leqslant x_{(n)}\alpha^{-1/n}\right\} = 1 - \alpha$$

и что он короче интервала, полученного в упражнении 20.2.

20.16 Используя результат упражнения 20.14, показать, что доверительные интервалы для θ в случае распределения

$$dF = e^{-(x-\theta)}\,dx, \qquad \theta \leqslant x \leqslant \infty,$$

получаются из соотношения

$$\mathbf{P}\left\{x_{(1)} + \frac{1}{n}\log\alpha \leqslant \theta \leqslant x_{(1)}\right\} = 1 - \alpha.$$

(Хузурбазар, 1955.)

20.17 Пусть $I(x)$ — доверительный интервал для параметра θ, $f(x|\theta)$ — плотность распределения и θ_0 — истинное значение параметра θ. Показать, что ожидаемая длина интервала $I(x)$ может быть записана как

$$\mathbf{M}(L) = \int\left\{\int\limits_{\theta \in I(x)} d\theta\right\} dF(x \mid \theta_0)$$

и что

$$\mathbf{M}(L) = \int\limits_{\theta \neq \theta_0} \mathbf{P}\{\theta \in I(x) \mid \theta_0\}\,d\theta$$

есть интеграл по всем ложным значениям θ от вероятности того, что θ попадает в доверительный интервал.

(Прэтт, 1961, 1963.)

20.18 Пусть в упражнении 20.17 $x \in A(\theta)$ тогда и только тогда, когда $\theta \in I(x)$. Показать, что $\mathbf{M}(L)$ минимизируется при выборе такого $A(\theta)$ при каждом θ, для которого $\mathbf{P}\{x \in A(\theta)\,|\,\theta_0\}$ минимальна. (Это эквивалентно выбору для каждого θ наиболее мощного критерия против альтернативного значения θ_0 (см. **23.26**, ниже).)

(Прэтт, 1961, 1963.)

20.19 Пусть x и y имеют двумерное нормальное распределение с дисперсиями σ_1^2, σ_2^2 и коэффициентом корреляции ρ. Показать, что величины

$$u = \frac{x}{\sigma_1} + \frac{y}{\sigma_2}, \qquad v = \frac{x}{\sigma_1} - \frac{y}{\sigma_2}$$

независимы и нормально распределены. Для выборки объема n с выборочными дисперсиями s_x^2 и s_y^2 и коэффициентом корреляции r_{xy} показать, что выборочный коэффициент корреляции между u и v может быть записан в виде

$$r_{uv}^2 = \frac{(l-\lambda)^2}{(l+\lambda)^2 - 4r_{xy}^2 l\lambda},$$

где $l = s_1^2/s_2^2$ и $\lambda = \sigma_1^2/\sigma_2^2$. Вывести отсюда, что, каково бы ни было значение ρ, доверительные границы для λ дается выражениями $l\{K - (K^2-1)^{1/2}\}$, $l\{K + (K^2-1)^{1/2}\}$, где $K = 1 + \dfrac{2(1-r_{xy}^2)}{n-2} t_\alpha^2$ и t_α^2 есть 100α-процентная точка t^2-распределения Стьюдента.

(Питмэн, 1939а.)

20.20 Показать, что величина $r(\bar{x}, \gamma)$, определенная в **20.39** соотношением (20.80), асимптотически по n равна

$$r(\bar{x}, \gamma) \sim r(0, \gamma)\left(1 + \frac{1}{2n}\right).$$

(Боукер, 1946.)

20.21 Используя метод из примера 6.4, показать, что для χ^2-распределения с ν степенями свободы значение, выше которого лежит 100β процентов распределения, есть χ_β^2, где

$$\frac{\chi_\beta^2}{\nu} \sim 1 + \left(\frac{2}{\nu}\right)^{1/2} d_{1-\beta} + \frac{2}{3\nu}\left(d_{1-\beta}^2 - 1\right) + o\left(\frac{1}{\nu}\right),$$

и

$$\int_{-\infty}^{d_\alpha} (2\pi)^{-1/2} \exp\left(-\frac{1}{2}t^2\right) dt = \alpha.$$

20.22 Комбинируя результаты упражнений 20.20, 20.21, показать, что согласно (20.83)

$$\lambda \sim r(0, \gamma)\left\{1 + \frac{d_\beta}{(2n)^{1/2}} + \frac{5\left(d_\beta^2 + 2\right)}{12n}\right\}.$$

(Боукер, 1946.)

ГЛАВА 21

ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ: ФИДУЦИАЛЬНЫЕ ИНТЕРВАЛЫ

21.1 В начале этой главы сделаем небольшое замечание о терминологии. Задачи интервального оценивания (в его точном смысле) начали привлекать внимание статистиков примерно в 1925—1930 гг. Подход, основанный на доверительных интервалах, которые были определены в предыдущей главе, и подход, основанный на фидуциальных интервалах, к изложению которых мы переходим, были предложены соответственно Дж. Нейманом и Р. А. Фишером. Так как оба подхода давали, казалось, одинаковые результаты, то вначале было естественное убеждение, что оба эти метода выражают одно и то же, но в разных терминах. Как следствие этого, в ранней литературе часто употреблялось название «фидуциальные» интервалы в смысле наших «доверительных» и (менее часто) «доверительные» интервалы в некотором смысле, более близком к «фидуциальному» методу рассуждений.

Хотя путаница с названиями никогда не была до конца прояснена, сейчас общепризнано, что фидуциальные интервалы отличаются по своей природе от доверительных. Но сторонники этих подходов, как нам кажется, не всегда с полной ясностью указывают, в чем состоит различие; также не всегда они используют термин «фидуциальный» в строгом соответствии с тем значением, в котором его употреблял Фишер, первым предложивший его. Мы изложим основные, на наш взгляд, идеи фидуциального подхода, но читатель, обращающийся к первоисточникам, должен иметь в виду, что он может встретить упомянутое несовпадение в терминологии.

21.2 Для того чтобы понять основные идеи, обратимся к примеру. Рассмотрим выборку объема n из нормальной совокупности с неизвестным средним μ и единичной дисперсией. Достаточной статистикой для μ является выборочное среднее $\bar{x}$, которое имеет распределение

$$dF = \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{1}{2} n (\bar{x} - \mu)^2\right\} d\bar{x}. \qquad (21.1$$

Выражение (21.1) есть распределение различных значений $\bar{x}$ при фиксированном неизвестном μ. Теперь предположим, что

мы имеем единственную выборку объема n с выборочным средним $\bar{x}_1$. Согласно (17.68) функция правдоподобия этой выборки $L(\bar{x}_1|\mu)$ будет зависеть от μ только через распределение $\bar{x}$ (21.1) (так как $\bar{x}$ — достаточная статистика для μ), и поэтому последнее может быть взято в качестве функции правдоподобия. Итак,

$$L(\bar{x}_1|\mu) \propto \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{1}{2} n(\bar{x}_1 - \mu)^2\right\}. \tag{21.2}$$

Если мы готовы, возможно несколько интуитивно, измерять функцией правдоподобия (21.2) степень нашего доверия к конкретному значению μ, то окончательно можно написать

$$dF = \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{1}{2} n(\bar{x}_1 - \mu)^2\right\} d\mu. \tag{21.3}$$

Последнее распределение будем называть *фидуциальным распределением* параметра μ. Интеграл от выражения (21.3) по прямой $(-\infty, \infty)$ равен 1, так что нормирующей константы здесь не требуется.

21.3 Фидуциальное распределение не является частотным распределением в том смысле, в каком использовалось это выражение до сих пор. Это новое понятие, выражающее интенсивность нашей веры в различные возможные значения параметра. Нередко бывает, что недифференциальные элементы, как в случае (21.3) и (21.1), совпадают, но это обстоятельство несущественно.

Фидуциальное распределение не является также распределением вероятностей в смысле частотной теории вероятностей. Его можно рассматривать как распределение вероятностей в смысле степеней убежденности; ниже будет рассмотрена связь с интервальным оцениванием, основанным на использовании теоремы Байеса. Фидуциальное распределение можно также рассматривать как новое понятие, дающее формальное выражение некоторым интуитивным идеям относительно величины нашего доверия к различным значениям параметра μ.

21.4 Фидуциальное распределение может быть теперь использовано для определения интервалов, внутри которых заключено μ. Выберем числа, скажем 0,02275 и 0,97725, и будем рассматривать их как критические в том смысле, что при любом приемлемом значении μ наблюденному значению $\bar{x}$ должна соответствовать (кумулятивная) вероятность не меньшая, чем 0,02275, и не большая, чем 0,97725. Так как этим значениям соответствуют уклонения $\pm 2\sigma$ от среднего значения нормальной совокупности, и $\sigma = 1/\sqrt{n}$, то имеем

$$-2 \leqslant (\bar{x}_1 - \mu)\sqrt{n} \leqslant 2,$$

что эквивалентно

$$\bar{x}_1 - 2/\sqrt{n} \leqslant \mu \leqslant \bar{x}_1 + 2/\sqrt{n}. \qquad (21.4)$$

Случилось так, что мы получили то же самое неравенство, к которому нас привел в примере 20.1 центральный доверительный интервал, основанный на распределении (21.1). Существенно заметить, что этот вывод достигается другим ходом рассуждений. Доверительный подход говорит, что если мы утверждаем (21.4), то будем правы в 95,45% случаев при многократных выборках. При фидуциальном подходе утверждение (21.4) эквивалентно тому, что (в некотором смысле, не имеющем точного определения) мы на 95,45% уверены в том, что окажемся правы в данном конкретном случае. Такое же смещение акцента нам встречалось при рассмотрении самой функции правдоподобия: функция $L(x|\theta)$ может рассматриваться или как элементарная вероятность, когда параметр θ фиксирован, а x меняется, или как величина правдоподобия, когда x фиксировано, а θ меняется.

Так же и здесь мы можем делать вывод о параметре θ, считая его константой и строя содержащие его интервалы, которые являются случайными величинами, либо считая фиксированными наблюдения и строя интервалы, основанные на некоторой неопределяемой степени доверия к значениям параметра, порождающего эти наблюдения.

21.5 Существует еще одно существенное различие между этими двумя методами. В предыдущей главе мы видели, что в доверительной теории возможны различные множества доверительных интервалов для одного и того же параметра, основанные на различных статистиках (хотя мы, естественно, различаем эти множества и выбираем наиболее короткие или наиболее селективные интервалы). Ничего подобного нет в фидуциальной теории (даже в смысле выбора центральных или нецентральных интервалов для одного и того же распределения, когда используются оба его хвоста). Мы должны использовать всю информацию о параметре, которая содержится в функции правдоподобия. Отсюда следует, что если пределы для параметра θ определяются единственной статистикой t, то она должна быть достаточной для θ. (К этому же заключению мы пришли с точки зрения наиболее селективных интервалов в **20.30**.)

Как указывалось в **17.38**, всегда существует система совместно достаточных статистик для одного неизвестного параметра, а именно сами n наблюдений. Но эта тавтология малоутешительна, так как трудно использовать даже достаточные системы, состоящие из двух статистик, а более широкие системы практически бесполезны. Вопрос о том, как строить интервал для

одного параметра, когда не существует одномерной достаточной статистики, фидуциальная теория по большей части обходит молчанием.

21.6 Пусть $f(t, \theta)$ — непрерывная плотность, а $F(t, \theta)$ — функция распределения статистики t, являющейся достаточной для параметра θ. Рассмотрим поведение функции f при изменении θ, когда фиксировано некоторое значение t. Предположим, что нам заранее известен интервал изменения параметра θ, в частности это может быть интервал $(-\infty, \infty)$. Выберем некоторую критическую вероятность $1-\alpha$ (аналог коэффициента доверия), и пусть θ_α — значение параметра θ такое, что $F(t, \theta) = 1-\alpha$.

Предположим теперь, что при любом t_1 функция $f(t_1, \theta)$ является монотонно невозрастающей в интервале изменения параметра θ. Тогда при всех $\theta \leqslant \theta_\alpha$ значение плотности вероятности в точке t_1 не меньше, чем $f(t_1, \theta_\alpha)$, а при $\theta > \theta_\alpha$ — меньше. В этом случае мы выбираем интервал $\theta \leqslant \theta_\alpha$ в качестве фидуциального интервала. Он содержит в себе все те значения параметра, при которых плотность вероятности больше или равна $f(t_1, \theta_\alpha)$.

21.7 Если нам требуется фидуциальный интервал вида $\theta_{\alpha_1} < \theta \leqslant \theta_{\alpha_2}$, то мы должны найти такие два значения параметра θ, что $f(t_1, \theta_{\alpha_1}) = f(t_1, \theta_{\alpha_2})$ и $F(t_1, \theta_{\alpha_2}) - F(t_1, \theta_{\alpha_1}) = 1-\alpha$. И если при θ, лежащем между этими значениями, $f(t_1, \theta)$ больше, чем $f(t_1, \theta_{\alpha_1})$ или $f(t_1, \theta_{\alpha_2})$, а при θ, лежащем вне этого интервала, $f(t_1, \theta)$ меньше этой величины, то интервал вновь содержит те значения параметра, при которых функция плотности не меньше, чем значения плотности в критических точках.

Если распределение статистики t симметрично, то по обе стороны интервала остаются одинаковые по площади хвосты. При несимметричном распределении эти хвосты должны в сумме содержать вероятность α, но вероятности хвостов не равны между собой. Равными должны быть значения плотности в концах интервала. Подобные вопросы уже обсуждались нами в связи с центральными доверительными интервалами в **20.7**.

21.8 При таком условии, если мы увеличим фидуциальный интервал с каждой стороны на величину $d\theta$, то значение функции распределения в концевой точке уменьшается на $(\partial F(t_1, \theta)/\partial\theta)\, d\theta$. Следовательно, для фидуциального распределения мы имеем

$$dF = -\frac{\partial F(t_1, \theta)}{\partial \theta}\, d\theta. \qquad (2.5)$$

Эта формула, однако, требует, чтобы $f(t_1, \theta)$ была неубывающей функцией от θ в нижнем конце интервала и невозрастающей в верхнем.

Пример 21.1

Рассмотрим вновь нормальное распределение из (21.1). Для любого фиксированного $\bar{x}_1$ при изменении μ от $-\infty$ через $\bar{x}_1$ до $+\infty$ плотность вероятности возрастает монотонно от нуля до максимума в $\bar{x}_1$ и затем убывает монотонно до нуля. Таким образом, значение плотности для любого μ в интервале от $\bar{x}_1 - k$ до $\bar{x}_1 + k$ больше, чем ее значения в точках $\bar{x}_1 - k$ или $\bar{x}_1 + k$. Следовательно, в качестве фидуциального интервала мы можем взять интервал $\bar{x}_1 - k \leqslant \mu \leqslant \bar{x}_1 + k$ для любого подходящего значения $k > 0$. В (21.4) мы выбрали k равным $2/\sqrt{n}$.

Пример 21.2

В качестве примера несимметричного выборочного распределения рассмотрим распределение

$$dF = \frac{x^{p-1}e^{-x/\theta}}{\theta^p \Gamma(p)}\,dx, \qquad p > 0, \qquad 0 \leqslant x \leqslant \infty. \tag{21.6}$$

Если p известно, то $t \equiv \bar{x}/p$ является достаточной статистикой для θ (см. упражнение 17.1), и легко видеть, что ее выборочное распределение есть

$$dF = \left(\frac{\beta}{\theta}\right)^{\beta} \frac{t^{\beta-1}e^{-\beta t/\theta}}{\Gamma(\beta)}\,dt, \tag{21.7}$$

где $\beta = np$. В данном случае параметр θ может изменяться только от 0 до ∞. При этом функция в (21.7) при фиксированном t монотонно возрастает от 0 до максимума и затем убывает вновь до 0, представляя собой в действительности обращенное распределение типа III. Таким образом, если мы определим θ_{α_1} и θ_{α_2} так, чтобы значения функции в этих точках совпадали, а интеграл от распределения (21.7) в этих пределах равнялся заданному значению $1 - \alpha$, то фидуциальным интервалом будет $\theta_{\alpha_1} \leqslant \theta \leqslant \theta_{\alpha_2}$.

Выражение (21.7) можно записать в виде

$$dF = \left(\frac{\beta t}{\theta}\right)^{\beta-1} \frac{e^{-\beta t/\theta}}{\Gamma(\beta)}\,d\left(\frac{\beta t}{\theta}\right), \tag{21.8}$$

и, следовательно,

$$F(t, \theta) = \int_0^{\beta t/\theta} \frac{u^{\beta-1}e^{-u}}{\Gamma(\beta)}\,du. \tag{21.9}$$

Отсюда

$$-\frac{\partial F}{\partial \theta} = -\left[\frac{u^{\beta-1}e^{-u}}{\Gamma(\beta)}\right]_{u=\beta t/\theta} \frac{\partial}{\partial \theta}\left(\frac{\beta t}{\theta}\right) = \left(\frac{\beta t}{\theta}\right)^{\beta-1} \frac{e^{-\beta t/\theta}}{\Gamma(\beta)} \frac{\beta t}{\theta^2}.$$

Таким образом, фидуциальное распределение для θ есть

$$\left(\frac{\beta t}{\theta}\right)^{\beta} \frac{e^{-\beta t/\theta}}{\Gamma(\beta)} \frac{d\theta}{\theta}. \qquad (21.10)$$

Интеграл от этого выражения в пределах от $\theta = 0$ до $\theta = \infty$ равен единице.

Сравнивая (21.7) с (21.10), замечаем, что мы заменили dt не на $d\theta$, а на $t\,d\theta/\theta$. Можно сказать несколько иначе, что мы заменили dt/t на $d\theta/\theta$. Полезно рассмотреть, почему это так происходит, и повторить в специфической форме рассуждения из **21.8**.

Мы определяем наш фидуциальный интервал, используя вероятность $F(t, \theta)$. Глядя на (21.9), мы видим, что это есть интеграл, верхний предел которого, с точностью до постоянной, равен t/θ. Поэтому при дифференцировании по θ подинтегральная функция умножается на $d_\theta(t/\theta) = -t\,d\theta/\theta^2$, в то время как при дифференцировании по t она умножается на $d_t(t/\theta) = dt/\theta$. Таким образом, согласно (21.5) $-(\partial F/\partial\theta)\,d\theta = t\,d\theta/\theta^2$, тогда как $(\partial F/\partial t)\,dt = dt/\theta$. Приравнивая левые и правые части этих выражений, получаем $d\theta/\theta = dt/t$.

21.9 Попытка обобщить нашу теорию на случай нескольких параметров приводит к трудностям. Дело в том, что практических задач в этой области так мало, что любая общая теория повисает в воздухе из-за недостатка иллюстрирующих ее примеров. Поэтому мы рассмотрим лишь два важных стандартных случая: оценку среднего в нормальной выборке при неизвестной дисперсии и оценку разности двух средних по выборкам из двух нормальных совокупностей с неравными дисперсиями.

Фидуциальные выводы для распределения Стьюдента

21.10 Известно, что в нормальных выборках выборочное среднее $\bar{x}$ и выборочная дисперсия $s^2 \left(= \sum (x - \bar{x})^2/n\right)$ являются совместно достаточными статистиками для среднего μ и дисперсии σ^2. Их распределение имеет вид

$$dF \propto \frac{1}{\sigma} \exp\left\{-\frac{n}{2\sigma^2}(\bar{x} - \mu)^2\right\} d\bar{x} \left(\frac{s}{\sigma}\right)^{n-2} \exp\left\{-\frac{ns^2}{2\sigma^2}\right\} \frac{ds}{\sigma}. \qquad (21.11)$$

Если бы мы строили фидуциальные границы для μ при известном σ, то использовали бы первый множитель в правой части (21.11); при построении же границ для σ, когда μ известно, нельзя использовать только второй множитель, так как σ входит и в первый множитель. На самом деле (см. пример 17.10) в этом случае достаточной статистикой является не s^2, а величина $\sum (x - \mu)^2/n$; чтобы получить ее распределение, требуются оба множителя в (21.11).

Для получения фидуциального распределения параметра μ при известном σ следует, как и в примере 21.1, заменить $d\bar{x}$ на $d\mu$. При известном же μ нужно воспользоваться тем, что величина $\sum(x-\mu)^2 = s'^2$ распределена, как x в (21.6) с $p=n$ и $\theta=\sigma^2$, и поэтому, как в примере 21.2, заменить ds'/s' на $d\sigma/\sigma$. В соотношении (21.11) статистика s распределена, как s', но с $p=n-1$. Возникает вопрос: можем ли мы заменить $d\bar{x}\,ds/s$ в (21.11) на $d\mu\,d\sigma/\sigma$, чтобы получить совместное фидуциальное распределение параметров μ и σ?

Фидуциалисты считают, что это можно сделать. Нам этот вопрос представляется очень спорным *). Однако сделаем такое допущение и посмотрим, куда оно нас приведет. В этом случае в качестве фидуциального распределения мы имеем

$$dF \propto \frac{1}{\sigma}\exp\left\{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2\right\}d\mu\left(\frac{s}{\sigma}\right)^{n-1}\exp\left\{-\frac{ns^2}{2\sigma^2}\right\}\frac{d\sigma}{\sigma}. \quad (21.12)$$

Проинтегрировав его теперь по σ для того, чтобы получить фидуциальное распределение параметра μ, приходим к выражению

$$dF \propto \frac{d\mu/s}{\left\{1+\frac{(\bar{x}-\mu)^2}{s^2}\right\}^{n/2}}. \quad (21.13)$$

Это есть распределение Стьюдента с $\frac{\mu-\bar{x}}{s}\sqrt{n-1}$ вместо обычного t и с $n-1$ степенями свободы. Таким образом, при заданном α мы можем найти два значения t, t_0 и t_1, для которых

$$\mathsf{P}\{-t_0 \leqslant t \leqslant t_1\} = 1-\alpha.$$

Последнее эквивалентно тому, что параметр μ заключен в интервале

$$\{\bar{x}-st_0/\sqrt{n-1},\ \bar{x}+st_1/\sqrt{n-1}\}. \quad (21.14)$$

Полученное соотношение может быть интерпретировано, как и в **20.31**, в терминах доверительных интервалов, а именно, если мы утверждаем, что μ лежит в интервале (21.14), то будем правы в $1-\alpha$-й части случаев. Но, как мы увидим позже, это отнюдь не существенно для фидуциального подхода.

*) Хотя $\bar{x}$ и s являются статистически независимыми, μ и σ не являются независимыми ни в каком фидуциальном смысле. Законы перехода от частотных распределений к фидуциальным совершенно не выяснены в случае многих параметров. В приведенном выше примере некоторое оправдание этому переходу можно вывести *a posteriori* из того, что он приводит к распределению Стьюдента, но если фидуциальную теорию принимать самое по себе, то нужно требовать что-то дополнительно.

Задача о двух средних

21.11 Мы теперь переходим к задаче нахождения интервальной оценки для разности между средними двух нормальных совокупностей. Эту задачу мы оставили без рассмотрения в предыдущей главе для того, чтобы иметь возможность дать здесь единое изложение. Мы обсудим сначала различные подходы к задаче с позиций доверительных интервалов и затем перейдем к ее решению с помощью фидуциальных интервалов. В заключение рассмотрим задачу с точки зрения теоремы Байеса.

21.12 Предположим, что у нас есть два нормальных распределения с параметрами соответственно μ_1, σ_1^2 и μ_2, σ_2^2, из которых извлечены выборки объемов n_1, n_2 соответственно, и выборочные средние и дисперсии равны $\bar{x}_1$, s_1^2 и $\bar{x}_2$, s_2^2. Без потери общности предположим, что $n_1 \leqslant n_2$.

В случае, когда $\sigma_1^2 = \sigma_2^2 = \sigma^2$, задача нахождения интервала для $\mu_1 - \mu_2 = \delta$ проста. В этом случае величина $d = \bar{x}_1 - \bar{x}_2$ имеет нормальное распределение с

$$\mathbf{M}(d) = \delta, \qquad \mathbf{D}d = \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \tag{21.15}$$

и каждая из величин $n_1 s_1^2/\sigma_1^2$, $n_2 s_2^2/\sigma_2^2$ имеет χ^2-распределение с $n_1 - 1$ и $n_2 - 1$ степенями свободы соответственно. Поскольку две выборки независимы, величина $(n_1 s_1^2 + n_2 s_2^2)/\sigma^2$ будет распределена как χ^2 с $n_1 + n_2 - 2$ степенями свободы, и, полагая

$$s^2 = (n_1 s_1^2 + n_2 s_2^2)/(n_1 + n_2 - 2),$$

имеем

$$\mathbf{M}(s^2) = \sigma^2. \tag{21.16}$$

Теперь статистика

$$y = \frac{d - \delta}{\left\{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right\}^{1/2}} \Bigg/ \left(\frac{s^2}{\sigma^2}\right)^{1/2} = \tag{21.17}$$

$$= \frac{d - \delta}{\left\{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right\}^{1/2}} \tag{21.18}$$

есть отношение нормированной нормальной величины к корню квадратному из несмещенной оценки для дисперсии. При этом знаменатель не зависит от числителя (так как s_1^2 и s_2^2 не зависят от $\bar{x}_1$ и $\bar{x}_2$). Кроме того, величина $(n_1 + n_2 - 2)s^2/\sigma^2$ распределена как χ^2 с $n_1 + n_2 - 2$ степенями свободы (ст. св.).

Таким образом, y имеет такой же вид, как и отношение (при одной выборке)

$$\frac{\bar{x}_1-\mu_1}{\{s_1^2/(n_1-1)\}^{1/2}}=\frac{\bar{x}_1-\mu_1}{(\sigma^2/n_1)^{1/2}}\Bigg/\left\{\frac{n_1 s_1^2/(n_1-1)}{\sigma^2}\right\}^{1/2},$$

которое нам уже встречалось несколько раз (см. пример 11.8) и которое имеет распределение Стьюдента с n_1-1 ст. св. Следовательно, статистика (21.18) имеет также распределение Стьюдента, но с n_1+n_2-2 ст. св. Этот результат можно было доказать и непосредственно.

В этом случае построение доверительных или фидуциальных интервалов не вызывает затруднений: мы просто используем метод из **20.31** или из **21.10**; конечно, как и в случае одной выборки, получатся одинаковые результаты.

21.13 Трудности возникают, когда мы переходим к случаю $\sigma_1^2\neq\sigma_2^2$. Случайная величина, имеющая распределение Стьюдента с n_1+n_2-2 ст. св. (аналогичная (21.17)), теперь имеет вид

$$t=\frac{d-\delta}{\left\{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}\right\}^{1/2}}\Bigg/\left\{\frac{\dfrac{n_1 s_1^2}{\sigma_1^2}+\dfrac{n_2 s_2^2}{\sigma_2^2}}{n_1+n_2-2}\right\}^{1/2}. \qquad (21.19)$$

Числитель в (21.19) есть нормированная нормальная величина, а знаменатель является корнем квадратным из независимой от числителя случайной величины, имеющей распределение χ^2, деленной на число степеней свободы (как и в (21.17)). Сложность состоит в том, что в (21.19) входит неизвестное отношение дисперсий $\theta=\sigma_1^2/\sigma_2^2$. Если положить $u=s_1^2/s_2^2$, $N=n_1/n_2$, то (21.19) можно переписать так:

$$t=\frac{(d-\delta)(n_1+n_2-2)^{1/2}}{s_2\left\{\left(1+\dfrac{\theta}{N}\right)\left(1+\dfrac{Nu}{\theta}\right)\right\}^{1/2}}, \qquad (21.20)$$

что ясно указывает на зависимость от неизвестного параметра θ. Конечно, при $\theta=1$ (21.20) переходит в (21.18).

21.14 Рассмотрим теперь методы, с помощью которых можно устранить «мешающий параметр» θ из интервальных утверждений, касающихся δ. Мы должны, естественно, найти некоторую статистику, отличную от t в (21.20). Одна возможность

напрашивается сама собой при рассмотрении второй записи, (21.18), для выражения (21.17) при $\theta = 1$. Статистика

$$z = \frac{d - \delta}{\left(\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}\right)^{1/2}} \tag{21.21}$$

является, подобно (21.18), отношением нормальной величины с нулевым средним к корню квадратному из независимой от нее несмещенной оценки дисперсии числителя. Однако эта оценка имеет не χ^2-распределение и, следовательно, z не есть величина Стьюдента. Как мы увидим ниже, на статистике z основан фидуциальный подход к этой задаче и один подход, дающий приближенный доверительный интервал.

Другая возможность состоит в исследовании распределения самой статистики (21.18) с тем, чтобы увидеть, до какой степени статистика, пригодная для случая $\theta = 1$, сохраняет свои свойства при $\theta \neq 1$. Это исследование было также проведено с точки зрения доверительных интервалов.

Однако прежде чем перейти к обсуждению подходов, намеченных в этом пункте, мы подробно рассмотрим точное доверительно-интервальное решение задачи, основанное на распределении Стьюдента и его свойствах. Эти результаты принадлежат Шеффе (1943а, 1944).

Точные доверительные интервалы, основанные на распределении Стьюдента

21.15 Если мы хотим построить для параметра δ точный доверительный интервал, основанный на распределении Стьюдента, то достаточно найти линейную и квадратичную функции от наблюдений L и Q, соответственно, такие, что для любых значений σ_1^2, σ_2^2

(I) L и Q независимы,
(II) $\mathsf{M}(L) = \delta$ и $\mathsf{D}L = V$,
(III) Q/V имеет χ^2-распределение с k ст. св.

Тогда величина

$$t = \frac{L - \delta}{(Q/k)^{1/2}} \tag{21.22}$$

подчиняется распределению Стьюдента с k ст. св. Сейчас мы докажем замечательный результат, принадлежащий Шеффе (1944), о том, что не существует статистики вида (21.22), являющейся симметричной функцией от наблюдений каждой выборки. Иначе говоря, t не может быть инвариантной относительно перестановки элементов первой выборки x_{1i}

($i = 1, 2, \ldots, n_1$) между собой и элементов второй выборки x_{2i} ($i = 1, 2, \ldots, n_2$) между собой.

21.16 Предположим, что статистика t симметрична в указанном смысле. Тогда мы должны иметь

$$\left.\begin{aligned} L &= c_1 \sum_i x_{1i} + c_2 \sum_i x_{2i}, \\ Q &= c_3 \sum x_{1i}^2 + c_4 \sum_{i \neq j} x_{1i} x_{1j} + c_5 \sum x_{2i}^2 + \\ &\qquad + c_6 \sum_{i \neq j} x_{2i} x_{2j} + c_7 \sum_{i,j} x_{1i} x_{2j}, \end{aligned}\right\} \tag{21.23}$$

где коэффициенты c_i являются постоянными, не зависящими от параметров. Из (21.22) вытекает, что

$$\mathbf{M}(L) = \delta = \mu_1 - \mu_2, \tag{21.24}$$

а из (21.23), что

$$\mathbf{M}(L) = c_1 n_1 \mu_1 + c_2 n_2 \mu_2. \tag{21.25}$$

Выражения (21.24) и (21.25) являются тождествами относительно μ_1 и μ_2; следовательно,

$$c_1 n_1 \mu_1 = \mu_1, \qquad c_2 n_2 \mu_2 = -\mu_2,$$

поэтому

$$c_1 = 1/n_1, \qquad c_2 = -1/n_2. \tag{21.26}$$

Из (21.26) и (21.23) получаем

$$L = \bar{x}_1 - \bar{x}_2 = d \tag{21.27}$$

и, следовательно,

$$\mathbf{D}L = V = \sigma_1^2/n_1 + \sigma_2^2/n_2. \tag{21.28}$$

Поскольку отношение Q/V имеет χ^2-распределение с k ст. св., то

$$\mathbf{M}(Q/V) = k,$$

откуда, используя (21.28), находим

$$\mathbf{M}(Q) = k(\sigma_1^2/n_1 + \sigma_2^2/n_2), \tag{21.29}$$

тогда как из (21.23) вытекает, что

$$\begin{aligned} \mathbf{M}(Q) = {} & c_3 n_1 (\sigma_1^2 + \mu_1^2) + c_4 n_1 (n_1 - 1)\mu_1^2 + \\ & + c_5 n_2 (\sigma_2^2 + \mu_2^2) + c_6 n_2 (n_2 - 1)\mu_2^2 + c_7 n_1 n_2 \mu_1 \mu_2. \end{aligned} \tag{21.30}$$

Приравнивая (21.29) и (21.30), находим выражения для постоянных c_i. Используя их в (21.23), получаем

$$Q = k\left\{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}\right\}. \tag{21.31}$$

Соотношения (21.27) и (21.31) переводят (21.22) в (21.21). Далее, $n_1 s_1^2/\sigma_1^2$ и $n_2 s_2^2/\sigma_2^2$ — независимые χ^2-случайные величины, а линейная функция двух таких случайных величин может сама иметь χ^2-распределение лишь в том случае, если она является их суммой. Таким образом, из (21.31) следует, что Q тогда только будет иметь χ^2-распределение, когда

$$\frac{k\sigma_1^2}{n_1(n_1-1)} = \frac{k\sigma_2^2}{n_2(n_2-1)} = 1$$

или

$$\theta = \frac{\sigma_1^2}{\sigma_2^2} = \frac{n_1(n_1-1)}{n_2(n_2-1)}. \qquad (21.32)$$

При заданных n_1, n_2 последнее имеет место только для специальных значений σ_1^2, σ_2^2. Так как мы требуем, чтобы наше утверждение было справедливо для *всех* значений σ_1^2, σ_2^2, то мы пришли к противоречию. Итак, t не может быть симметричной функцией в указанном смысле.

21.17 Поскольку мы не можем найти симметричную функцию желаемого вида, имеющую распределение Стьюдента, то обратимся теперь к другим функциям. Рассмотрим (21.22) в случае, когда

$$\left.\begin{aligned} L &= \sum_{i=1}^{n_1} d_i/n_1, \\ Q &= \frac{1}{n_1}\sum_{i=1}^{n_1}(d_i - L)^2, \end{aligned}\right\} \qquad (21.33)$$

а d_i — одинаково распределенные нормальные величины с

$$\mathsf{M}(d_i) = \delta, \qquad \mathsf{D}d_i = \sigma^2. \qquad (21.34)$$

Напомним, что мы считаем $n_1 \leqslant n_2$. Теперь (21.22) превращается в величину

$$t = \frac{L-\delta}{\{Q/(n_1-1)\}^{1/2}} = (L-\delta)\left\{\frac{n_1(n_1-1)}{\sum(d_i-L)^2}\right\}^{1/2}, \qquad (21.35)$$

которая имеет распределение Стьюдента с $n_1 - 1$ ст. св.

Предположим, что в терминах исходных наблюдений

$$d_i = x_{1i} - \sum_{j=1}^{n_2} c_{ij} x_{2j}. \qquad (21.36)$$

Величины d_i, являясь линейными функциями от нормальных случайных величин (см. **15.4**), имеют многомерное нормальное

распределение. Для выполнения соотношений (21.34) необходимо и достаточно, чтобы

$$\left.\begin{aligned} \sum_j c_{ij} &= 1, \\ \sum_j c_{ij}^2 &= c^2, \\ \sum_j c_{ij}c_{kj} &= 0, \quad i \neq k. \end{aligned}\right\} \tag{21.37}$$

Таким образом, из (21.36) и (21.37) следует

$$\mathsf{D}(d_i) = \sigma^2 = \sigma_1^2 + c^2\sigma_2^2. \tag{21.38}$$

21.18 Центральный доверительный интервал, полученный на основе статистики (21.35) при коэффициенте доверия $1-\alpha$, есть

$$|L - \delta| \leqslant t_{n_1-1,\,\alpha}\{Q/(n_1-1)\}^{1/2}, \tag{21.39}$$

где $t_{n_1-1,\,\alpha}$ является соответствующим критическим значением при $n_1 - 1$ ст. св. Согласно (21.39) средняя длина интервала равна

$$\mathsf{M}(l) = 2t_{n_1-1,\,\alpha}\frac{\sigma}{\{n_1(n_1-1)\}^{1/2}}\,\mathsf{M}\left\{\left(\frac{n_1 Q}{\sigma^2}\right)^{1/2}\right\}, \tag{21.40}$$

где последний множитель в правой части можно найти, пользуясь тем, что величина n_1Q/σ^2 имеет χ^2-распределение с $n_1 - 1$ ст. св., а именно:

$$\mathsf{M}\left\{\left(\frac{nQ}{\sigma^2}\right)^{1/2}\right\} = \frac{\sqrt{2}\,\Gamma\left(\frac{1}{2}n_1\right)}{\Gamma\left\{\frac{1}{2}(n_1-1)\right\}}. \tag{21.41}$$

Для получения наименьшей средней длины в (21.40) мы должны минимизировать σ или, что эквивалентно, минимизировать c^2 в (21.38) при условии (21.37). Эту задачу можно представить геометрически следующим образом. Рассматривается пространство размерности n_2, причем каждому индексу j у c_{ij} соответствует координатная ось. Тогда уравнение $\sum_j c_{ij} = 1$ определяет гиперплоскость, а $\sum_j c_{ij}^2 = c^2$ определяет n_2-мерную гиперсферу, и они пересекаются по (n_2-1)-мерной гиперсфере. Требуется найти $n_1 \leqslant n_2$ векторов, выходящих из начала координат, концы которых лежат на этой (n_2-1)-мерной гиперсфере и которые (чтобы удовлетворить последнему из условий (21.37)) взаимно ортогональны, таким образом, чтобы минимизировать радиус n_2-мерной гиперсферы. Это можно сделать, взяв искомые векторы совпадающими с n_1 координатными

осями, и тогда $c^2 = 1$. Но если $n_1 < n_2$, то предложенную процедуру можно улучшить. Ниже будет показано, что мы можем, сохраняя ортогональность векторов, расположить их симметрично относительно прямой, составляющей равные углы с координатными осями, и тогда величина c^2 уменьшается от 1 до своего минимального значения, равного n_1/n_2.

21.19 Перепишем условия (21.37) в векторной форме:

$$\boldsymbol{c}_i\boldsymbol{u}' = 1, \quad \boldsymbol{c}_i\boldsymbol{c}_k' = \begin{cases} c^2, & i = k, \\ 0, & i \neq k, \end{cases} \tag{21.42}$$

где $\boldsymbol{c}_i$ есть i-я вектор-строка матрицы $\{c_{ij}\}$, а $\boldsymbol{u}$ — вектор-строка, состоящая из единиц.

Если n_1 векторов $\boldsymbol{c}_i$ удовлетворяют соотношениям (21.42), то можно присоединить к ним $n_2 - n_1$ векторов, удовлетворяющих второму условию (21.42) (нормировка и ортогональность), так, чтобы получить базис n_2-мерного пространства. Выразим $\boldsymbol{u}$ как линейную функцию от n_2 векторов $\boldsymbol{c}_i$:

$$\boldsymbol{u} = \sum_{k=1}^{n_2} g_k \boldsymbol{c}_k, \tag{21.43}$$

где g_k — скалярные величины. Теперь, используя (21.42) и (21.43), получаем

$$1 = \boldsymbol{c}_i\boldsymbol{u}' = \boldsymbol{c}_i \sum_{k=1}^{n_2} g_k \boldsymbol{c}_k' = \sum g_k \boldsymbol{c}_i\boldsymbol{c}_k' = g_i c^2.$$

Таким образом,

$$g_i = 1/c^2, \quad i = 1, 2, \ldots, n_1. \tag{21.44}$$

Поскольку $\boldsymbol{u}$ есть вектор-строка единиц, то

$$n_2 = \boldsymbol{u}\boldsymbol{u}' = \Big(\sum_k g_k \boldsymbol{c}_k\Big)\Big(\sum_k g_k \boldsymbol{c}_k'\Big).$$

Отсюда, учитывая (21.42), имеем

$$n_2 = \sum_{k=1}^{n_2} g_k^2 \boldsymbol{c}_k\boldsymbol{c}_k' = c^2\left(\sum_{k=1}^{n_1} + \sum_{n_1+1}^{n_2}\right) g_k^2. \tag{21.45}$$

Используя (21.44), из (21.45) находим

$$n_2 = c^2\left\{n_1/c^4 + \sum_{n_1+1}^{n_2} g_k^2\right\},$$

или

$$n_2 \geqslant \frac{n_1}{c^2}.$$

Следовательно,

$$c^2 \geqslant \frac{n_1}{n_2}, \tag{21.46}$$

что и требовалось доказать.

21.20 Знак равенства в (21.46) имеет место, когда $g_k = 0$ для $k = n_1 + 1, \ldots, n_2$. Тогда вектор $\boldsymbol{u}$ целиком лежит в пространстве, натянутом на первые n_1 $\boldsymbol{c}$-векторов. Из (21.44) видно, что эти векторы расположены симметрично вокруг него. Очевидно, существует бесконечное число способов определения c_{ij}, получаемых вращением набора из n_1 векторов. Шеффе (1943а) предложил особенно привлекательное решение:

$$\left.\begin{aligned} c_{jj} &= (n_1/n_2)^{1/2} - (n_1 n_2)^{-1/2} + 1/n_2, && j = 1, 2, \ldots, n_1, \\ c_{ij} &= -(n_1 n_2)^{-1/2} + 1/n_2, && j (\neq i) = 1, 2, \ldots, n_1, \\ c_{ij} &= 1/n_2, && j = n_1 + 1, \ldots, n_2. \end{aligned}\right\} \quad (21.47)$$

Легко проверить, что коэффициенты в (21.47) удовлетворяют условиям (21.37) с $c^2 = n_1/n_2$. Подстановка (21.47) в (21.36) дает

$$d_i = x_{1i} - (n_1/n_2)^{1/2} x_{2i} + (n_1 n_2)^{-1/2} \sum_{j=1}^{n_1} x_{2j} + (1/n_2) \sum_{j=1}^{n_2} x_{2j}. \quad (21.48)$$

Учитывая это, из (21.33) имеем

$$L = \bar{x}_1 - \bar{x}_2, \qquad Q = \frac{1}{n_1} \sum_{i=1}^{n_1} (u_i - \bar{u})^2, \quad (21.49)$$

где

$$u_i = x_{1i} - (n_1/n_2)^{1/2} x_{2i}, \qquad \bar{u} = \sum_{i=1}^{n_1} u_i/n_1. \quad (21.50)$$

Таким образом, из (21.35) и (21.48)—(21.50) получаем, что статистика

$$\{\bar{x}_1 - \bar{x}_2 - \delta\} \left\{ \frac{n_1 (n_1 - 1)}{\sum (u_i - \bar{u})^2} \right\}^{1/2} \quad (21.51)$$

имеет распределение Стьюдента с $n_1 - 1$ ст. св., и, используя ее, можно найти доверительные границы для $\delta = \mu_1 - \mu_2$.

21.21 Замечательно, что нам удалось точно решить задачу нахождения доверительного интервала, отказавшись только от естественного на вид требования симметрии. Свойство (21.51) имеет место для *любого* подмножества объема n_1, случайно выбранного из n_2 случайных величин, принадлежащих второй выборке. Подобно тому как в **20.22** мы прибегали к рандомизации для того, чтобы преодолеть трудности при определении точных доверительных интервалов для дискретной величины, так и здесь мы нашли, что одной лишь рандомизации достаточно, чтобы

устранить мешающий параметр θ. Но роль рандомизации не следует преувеличивать. Числитель в (21.51) содержит полностью выборочные средние обеих выборок, и только знаменатель изменяется в зависимости от случайного выбора подмножества во второй выборке. Интуитивно оценить, как много эффективности теряется при использовании этой процедуры, невозможно. Поэтому мы исследуем длину получаемых доверительных интервалов.

21.22 Из (21.38) и (21.46) для оптимального решения (21.48) имеем

$$\mathbf{D}d_i = \sigma^2 = \sigma_1^2 + (n_1/n_2)\sigma_2^2. \tag{21.52}$$

Подставляя (21.52) в (21.40) и используя (21.41), находим для средней длины доверительного интервала

$$\mathbf{M}(l) = 2t_{n_1-1,\,\alpha}\left\{\frac{\sigma_1^2 + (n_1/n_2)\sigma_2^2}{n_1(n_1-1)}\right\}^{1/2}\frac{\sqrt{2}\,\Gamma(n_1/2)}{\Gamma\{(n_1-1)/2\}}. \tag{21.53}$$

Сравним теперь этот интервал l с интервалом L, полученным из (21.19) при условии, что $\theta = \sigma_1^2/\sigma_2^2$ известно. Его средняя длина равна

$$\mathbf{M}(L) = 2t_{n_1+n_2-2,\,\alpha}\left\{\frac{\sigma_1^2/n_1 + \sigma_2^2/n_2}{n_1+n_2-2}\right\}^{1/2}\mathbf{M}\left\{\frac{n_1 s_1^2}{\sigma_1^2} + \frac{n_2 s_2^2}{\sigma_2^2}\right\}^{1/2}, \tag{21.54}$$

где последний множитель вычисляется с помощью χ^2-распределения с $n_1 + n_2 - 2$ ст. св. и равен

$$\frac{\sqrt{2}\,\Gamma\{(n_1+n_2-1)/2\}}{\Gamma\{(n_1+n_2-2)/2\}}. \tag{21.55}$$

Из (21.53)—(21.55) получаем отношение средних значений интервалов:

$$\mathbf{M}(l)/\mathbf{M}(L) = \\ = \frac{t_{n_1-1,\,\alpha}}{t_{n_1+n_2-2,\,\alpha}}\left(\frac{n_1+n_2-2}{n_1-1}\right)^{1/2}\frac{\Gamma(n_1/2)\,\Gamma\{(n_1+n_2-2)/2\}}{\Gamma\{(n_1-1)/2\}\,\Gamma\{(n_1+n_2-1)/2\}}. \tag{21.56}$$

Когда $n_1 \to \infty$ при фиксированном n_2, каждый из трех сомножителей в (21.56) стремится к 1 и, следовательно, отношение средних длин интервалов тоже стремится к единице, что интуитивно понятно. Если n_1 мало, то первые два сомножителя больше 1, а последний — меньше. Следующая таблица дает точные значения величины (21.56) для $1 - \alpha = 0{,}95$; 0,99 и нескольких объемов выборок.

Значения $\mathbf{M}(l)/\mathbf{M}(L)$ (из Шеффе, 1943а)

n_1-1 \ n_2-1	$1-\alpha=0{,}95$					$1-\alpha=0{,}99$				
	5	10	20	40	∞	5	10	20	40	∞
5	1,15	1,20	1,23	1,25	1,28	1,27	1,36	1,42	1,47	1,52
10		1,05	1,07	1,09	1,11		1,10	1,13	1,16	1,20
20			1,03	1,03	1,05			1,05	1,06	1,09
40				1,01	1,02				1,02	1,04
∞					1					1

Очевидно, что l является очень эффективным интервалом даже для выборок умеренных объемов. Его средняя длина превосходит среднюю длину L не более чем на 11% при $n_1 - 1 \geqslant 10$, $1-\alpha = 0{,}95$ и не более чем на 9% при $n_1 - 1 \geqslant 20$, $1-\alpha = 0{,}99$. Заметим, что мы сравнивали его с интервалом, *основанным на знании* θ. Принимая это во внимание, мы вполне можем сказать, что l действительно обладает очень хорошими качествами: рандомизация не внесла больших потерь эффективности.

В дополнение к приведенному мы рассмотрим приближенные доверительно-интервальные решения задачи о двух средних.

Приближенное нахождение доверительных интервалов

21.23 Уэлч (1938) исследовал приближенное распределение статистики (21.18), имеющей распределение Стьюдента при $\sigma_1^2 = \sigma_2^2$, в случае $\sigma_1^2 \neq \sigma_2^2$. В этом случае дисперсия числителя равна

$$\mathbf{D}(d-\delta) = \sigma_1^2/n_1 + \sigma_2^2/n_2,$$

и, полагая

$$\left.\begin{aligned} u &= (d-\delta)/(\sigma_1^2/n_1 + \sigma_2^2/n_2)^{1/2}, \\ w^2 &= s^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\Big/\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right), \end{aligned}\right\} \qquad (21.57)$$

выражение (21.18) можно переписать в виде

$$y = u/w. \qquad (21.58)$$

Трудность состоит в том, что, хотя w^2 и распределена независимо от u, она не является кратной χ^2 при $\theta \neq 1$. Однако, приравнивая ее первые два момента моментам χ^2-величины, мы можем определить число степеней свободы ν, при котором она

имеет *приближенно* χ^2-распределение. Согласно (21.57) ее среднее и дисперсия равны

$$\mathsf{M}(w^2) = b(\nu_1\theta + \nu_2), \qquad \mathsf{D}(w^2) = 2b^2(\nu_1\theta^2 + \nu_2), \tag{21.59}$$

где обозначено

$$\left.\begin{aligned} &\nu_1 = n_1 - 1, \quad \nu_2 = n_2 - 1, \\ &b = (n_1 + n_2)\sigma_2^2 / \{(n_1 + n_2 - 2)(n_2\sigma_1^2 + n_1\sigma_2^2)\}. \end{aligned}\right\} \tag{21.60}$$

Приравнивая (21.59) к моментам случайной величины χ^2 с ν ст. св., умноженной на постоянную g,

$$\mu_1' = g\nu, \qquad \mu_2 = 2g^2\nu, \tag{21.61}$$

находим

$$\left.\begin{aligned} &g = b(\theta^2\nu_1 + \nu_2)/(\theta\nu_1 + \nu_2), \\ &\nu = (\theta\nu_1 + \nu_2)^2/(\theta^2\nu_1 + \nu_2). \end{aligned}\right\} \tag{21.62}$$

При этих значениях g и ν w^2/g распределена приближенно как χ^2 с ν ст. св., и, следовательно, согласно (21.57) величина

$$t = u \Big/ \left\{ \frac{w^2}{g\nu} \right\}^{1/2} \tag{21.63}$$

имеет распределение Стьюдента с ν ст. св. Если $\theta = 1$, то $\nu = \nu_1 + \nu_2 = n_1 + n_2 - 2$, $g = b = 1/\nu$ и, как и должно быть, (21.63) сводится к (21.18). Но в общем случае g и ν зависят от θ.

21.24 Уэлч (1938) исследовал величину ошибки, к которой приводит предположение о том, что в соотношении (21.63) $\theta = 1$ (когда в действительности параметр принимает другие значения). Хотя его результаты сформулированы в терминах проверки гипотез, а не в терминах интервальных оценок, о которых у нас идет речь, мы их кратко упомянем. Он нашел, что при $n_1 = n_2$ нет большого вреда от незнания истинного значения параметра θ, но если $n_1 \neq n_2$, то могут возникнуть серьезные ошибки. Чтобы преодолеть эти трудности, он использовал в точности технику из **21.23** для аппроксимации распределения статистики z (21.21). Для этого случая им найдено, что, каковы бы ни были значения n_1 и n_2, статистика z имеет приближенно распределение Стьюдента с

$$\nu = \left(\frac{\theta}{n_1} + \frac{1}{n_2}\right)^2 \Big/ \left(\frac{\theta^2}{n_1^2(n_1 - 1)} + \frac{1}{n_2^2(n_2 - 1)}\right) \tag{21.64}$$

степенями свободы и что в этом случае влияние неверного предположения о параметре θ значительно меньше. Этого естественно было ожидать, поскольку знаменатель величины z в (21.21)

содержит отдельные оценки дисперсий σ_1^2, σ_2^2, тогда как в (21.58) входит совместная оценка s^2, которая менее пригодна, когда $\sigma_1^2 \neq \sigma_2^2$.

Лотон (1965) обобщил работу Я. Гаека, получив близкие верхние границы для размера и мощности симметричного двустороннего критерия, основанного на z.

Когда $\theta \to \infty$ или 0, то (21.64)$\to n_1 - 1$ или $n_2 - 1$ соответственно. Микки и Браун (1966) показали, что распределение статистики z ограничено распределениями Стьюдента с $n_1 + n_2 - 2$ и $\min(n_1 - 1, n_2 - 1)$ ст. св.

21.25 Уэлч (1947) усовершенствовал аппроксимативный подход предыдущего пункта. Его общие рассуждения в применении к нашей задаче можно вкратце изложить следующим образом. Определяя s_1^2, s_2^2 с делителями $n_1 - 1$, $n_2 - 1$ соответственно (чтобы это были несмещенные оценки дисперсий), мы ищем статистику $h(s_1^2, s_2^2, P)$ такую, что

$$\mathsf{P}\{(d - \delta) < h(s_1^2, s_2^2, P)\} = P, \tag{21.65}$$

каково бы ни было значение параметра θ. Теперь из того, что величина $(d - \delta)$ распределена нормально со средним нуль и дисперсией $\sigma_1^2/n_1 + \sigma_2^2/n_2 = D^2$ и не зависит от s_1^2, s_2^2, вытекает соотношение

$$\mathsf{P}\{(d - \delta) \leqslant h(s_1^2, s_2^2, P) \mid s_1^2, s_2^2\} = I(h/D), \tag{21.66}$$

где $I(x) = \int\limits_{-\infty}^{x} (2\pi)^{-1/2} \exp\left(-\frac{1}{2}t^2\right) dt$. Таким образом, из (21.65) и (21.66) получаем

$$P = \int\int I(h/D) f(s_1^2) f(s_2^2)\, ds_1^2\, ds_2^2. \tag{21.67}$$

Теперь можно разложить $I(h/P)$, как функцию от s_1^2, s_2^2, в ряд Тейлора в окрестности истинных значений σ_1^2, σ_2^2. Мы запишем это символически как

$$I\left\{\frac{h(s_1^2, s_2^2, P)}{D}\right\} = \exp\left\{\sum_{i=1}^{2} (s_i^2 - \sigma_i^2)\, \partial_i\right\} I\left\{\frac{h(s_1^2, s_2^2, P)}{s}\right\}, \tag{21.68}$$

где $s^2 = s_1^2/n_1 + s_2^2/n_2$, а оператор ∂_i обозначает дифференцирование по s_i^2 в точке $s_i^2 = \sigma_i^2$. Подставляя (21.68) в (21.67), имеем

$$P = \prod_{i=1}^{2} \left[\int \exp\{(s_i^2 - \sigma_i^2)\, \partial_i\} f(s_i^2)\, d(s_i^2)\right] I\left\{\frac{h(s_1^2, s_2^2, P)}{s}\right\}. \tag{21.69}$$

Далее, поскольку

$$f(s_i^2)\,ds_i^2 = \frac{1}{\Gamma(\nu_i/2)}\left(\frac{\nu_i s_i^2}{2\sigma_i^2}\right)^{\frac{1}{2}\nu_i - 1} \exp\left(-\frac{\nu_i s_i^2}{2\sigma_i^2}\right) d\left(\frac{\nu_i s_i^2}{\sigma_i^2}\right),$$

то, интегрируя в каждом из сомножителей выражения (21.69), находим

$$\int \exp\{(s_i^2 - \sigma_i^2)\,\partial_i\} f(s_i^2)\,ds_i^2 = \left(1 - \frac{2\sigma_i^2\partial_i}{\nu_i}\right)^{-\nu_i/2} \exp(-\sigma_i^2\partial_i),$$

что при подстановке в (21.69) дает

$$P = \prod_{i=1}^{2}\left[\left(1 - \frac{2\sigma_i^2\partial_i}{\nu_i}\right)^{-\nu_i/2} \exp(-\sigma_i^2\partial_i)\right] I\left\{\frac{h(s_1^2, s_2^2, P)}{s}\right\}. \quad (21.70)$$

Для любого заданного P мы можем решить уравнение (21.70) и получить вид функции h, т. е. найти $h(s_1^2, s_2^2, P)$.

Уэлч приводит разложение h в ряд, которое в нашем специальном случае имеет вид

$$\frac{h(s_1^2, s_2^2, P)}{s} = \xi\left[1 + \frac{(1+\xi)^2}{4}\sum_{i=1}^{2}\frac{c_i^2}{\nu_i} - \frac{1+\xi^2}{2}\sum_{i=1}^{2}\frac{c_i^2}{\nu_i^2} + \ldots\right], \quad (21.71)$$

где $c_i = \dfrac{s_i^2}{n_i}\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^{-1}$, $\nu_i = n_i - 1$ и ξ определяется из уравнения $I(\xi) = P$.

Согласно (21.21) $(d - \delta)/s = z$, поэтому соотношение (21.71) дает функцию распределения величины z.

Следуя работам Уэлча, Аспин (1948, 1949) и Трикетт и др. (1956) составили таблицы для (21.71) как функции от ν_1, ν_2 и c_1 при $P = 0{,}95$; 0,975; 0,99 и 0,995. Эти таблицы позволяют строить центральные доверительные границы для δ при $1 - \alpha = 0{,}90$; 0,95; 0,98 и 0,99. Некоторые из этих таблиц воспроизведены в таблице 11 *Biometrika Tables*.

Асимптотические выражения типа (21.71) были обоснованы Черновым (1949) и Д Уоллесом (1958). Разложение (21.71) является асимптотическим в том смысле, что каждый последующий член в правой части имеет более высокий порядок малости по ν_i, чем предыдущий.

Вальд (1955) значительно уточнил приближение Уэлча в случае $n_1 = n_2$.

Пресс (1966) показал, что для $1 - \alpha = 0{,}90$, $n_1 \leqslant n_2 \leqslant 30$ интервалы Уэлча имеют меньшую среднюю длину, чем (21.53), если θ мало (т. е. когда (21.51) отбрасывает информацию о бо-

лее изменчивой популяции), но это неверно при больших θ. Эти два множества интервалов асимптотически эквивалентны, и их средние длины никогда не отличаются более чем на 10% при $n_1 > 10$ и $0{,}01 \leqslant \theta \leqslant 100$.

Фидуциальное решение

21.26 Фидуциальное решение задачи о двух средних исходит из совместного распределения выборочных средних и дисперсий, которое можно представить в виде

$$dF \propto \frac{1}{\sigma_1\sigma_2} \exp\left\{-\frac{n_1}{2\sigma_1^2}(\bar{x}_1-\mu_1)^2 - \frac{n_2}{2\sigma_2^2}(\bar{x}_2-\mu_2)^2\right\} d\bar{x}_1\, d\bar{x}_2 \times$$

$$\times \frac{s_1^{n_1-2}}{\sigma_1^{n_1-2}} \frac{s_2^{n_2-2}}{\sigma_2^{n_2-2}} \exp\left\{-\frac{n_1}{2}\frac{s_1^2}{\sigma_1^2} - \frac{n_2}{2}\frac{s_2^2}{\sigma_2^2}\right\} ds_1\, ds_2. \quad (21.72)$$

Согласно фидуциальному подходу заменяем $d\bar{x}_1$, $d\bar{x}_2$ на $d\mu_1$, $d\mu_2$, а ds_1/s_1, ds_2/s_2 на $d\sigma_1/\sigma_1$, $d\sigma_2/\sigma_2$ (как в **21.10**). Тогда для фидуциального распределения (опуская степени s_1 и s_2, которые теперь являются постоянными) получаем

$$dF \propto \frac{1}{\sigma_1^{n_1+1}\sigma_2^{n_2+1}} \exp\left\{-\frac{n_1}{2\sigma_1^2}(\bar{x}_1-\mu_1)^2 - \frac{n_2}{2\sigma_2^2}(\bar{x}_2-\mu_2)^2\right\} d\mu_1\, d\mu_2 \times$$

$$\times \exp\left\{-\frac{n_1 s_1^2}{2\sigma_1^2} - \frac{n_2 s_2^2}{2\sigma_2^2}\right\} d\sigma_1\, d\sigma_2. \quad (21.73)$$

Обозначая

$$t_1 = \frac{(\mu_1-\bar{x}_1)\sqrt{n_1-1}}{s_1}, \qquad t_2 = \frac{(\mu_2-\bar{x}_2)\sqrt{n_2-1}}{s_2}, \quad (21.74)$$

находим, как в (21.10), совместное распределение параметров μ_1 и μ_2:

$$dF \propto \frac{d_\mu t_1}{\{1+t_1^2/(n_1-1)\}^{n_1/2}} \frac{d_\mu t_2}{\{1+t_2^2/(n_2-1)\}^{n_2/2}}, \quad (21.75)$$

где индекс μ у $d_\mu t_1$ указывает на то, что дифференциальный элемент равен $\sqrt{n_1-1}\, d\mu_1/s_1$ (и аналогично в случае второй выборки).

Мы не можем сразу перейти к нахождению интервала для $\delta = \mu_1 - \mu_2$. Действительно, из (21.74) имеем

$$(\mu_1-\bar{x}_1) - (\mu_2-\bar{x}_2) = \delta - d =$$

$$= t_1 s_1/\sqrt{n_1-1} - t_2 s_2/\sqrt{n_2-1}, \quad (21.76)$$

и, чтобы найти границы для δ, нам нужно знать фидуциальное распределение правой части (21.76) или некоторой подходящей

функции от нее. Эта правая часть является линейной функцией по t_1 и t_2, фидуциальное распределение которых дается соотношением (21.75). В данном случае Фишер (1935b, 1939), следуя Беренсу, выбрал в качестве наиболее удобной функции статистику (21.21):

$$z = \frac{d - \delta}{\left(\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}\right)^{1/2}}.$$

Имеем

$$z_1 = t_1 \cos\psi - t_2 \sin\psi, \tag{21.77}$$

где

$$\operatorname{tg}^2\psi = \frac{s_2^2}{n_2 - 1} \Big/ \frac{s_1^2}{n_1 - 1}. \tag{21.78}$$

При известном ψ распределение величины z (обычно называемое распределением Фишера — Беренса) может быть найдено из (21.76). Оно имеет не простой вид, но в *Statistical Tables* Фишера и Иэйтса приводятся критические значения z, когда заданы n_1, n_2, ψ и вероятность $1 - \alpha$. Используя эти таблицы (и вообще читая работы Фишера), нужно помнить, что наша величина $s^2/(n-1)$ обозначается у него как s'^2.

21.27 В этом случае, наиболее важном из рассмотренных до сих пор, фидуциальный подход не дает того же результата, как подход с доверительными интервалами. А именно, если мы определяем по вероятности $1 - \alpha$ соответствующие точки для z, скажем z_0 и z_1, и затем утверждаем, что

$$\bar{x}_1 - \bar{x}_2 - z_0 \sqrt{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}} \leqslant \mu_1 - \mu_2 \leqslant$$
$$\leqslant \bar{x}_1 - \bar{x}_2 + z_1 \sqrt{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}}, \tag{21.79}$$

то мы не окажемся правы в $1 - \alpha$-й части случаев при большом числе повторений. Это очевидно из того факта, что z представимо в виде

$$z = t \left\{ \frac{s_2^2 (1 + \theta/N)(1 + Nu/\theta)}{(n_2 - 1) s_1^2 + (n_1 - 1) s_2^2} \right\}^{1/2} \left\{ \frac{(n_1 - 1)(n_2 - 1)}{n_1 + n_2 - 2} \right\}^{1/2},$$

где t, определяемое соотношением (21.20), имеет точное распределение Стьюдента. Поскольку распределение величины t не зависит от θ, этого не может быть для статистики z.

Приведенный факт послужил последователям доверительно-интервального подхода основанием для критики. Фидуциалисты отвечают, что нет особых причин для того, чтобы такие

утверждения были справедливыми в некоторой части случаев при большом числе повторений, и что навязывать такое пожелание — значит упустить суть фидуциального подхода. Позднее мы вернемся к этому моменту.

Байесовские интервалы

21.28 Перейдем теперь к рассмотрению связи между фидуциальной теорией и интервальным оцениванием, основанным на теореме Байеса, следуя Джеффрейсу (1948). Эта теорема (**8.2—3**) утверждает, что вероятность события q_r при данных p и H пропорциональна произведению вероятности события q_r при заданном H на вероятность события p при фиксированных q_r и H. Или символически:

$$\mathbf{P}\{q_r \mid p,\ H\} \propto \mathbf{P}(q_r \mid H)\,\mathbf{P}(p \mid q_r,\ H). \qquad (21.80)$$

При байесовском подходе мы берем в качестве q_r значение оцениваемого параметра θ, а в качестве $\mathbf{P}(q_r \mid H)$ — его априорное распределение вероятностей. Тогда $\mathbf{P}(q_r \mid p, H)$ будет апостериорным распределением вероятностей параметра θ, и мы можем использовать его для установления границ, внутри которых лежит θ с заданной *в этом смысле* вероятностью.

Основной проблемой, как уже отмечалось ранее, является выбор априорного распределения $\mathbf{P}(q_r \mid H)$. Для того чтобы учесть различные ситуации, Джеффрейс расширил постулат Байеса (который устанавливает, что если ничего не известно о параметре θ и он изменяется в конечном интервале, то априорное распределение следует взять пропорциональным $d\theta$). В частности, (1) если интервал изменения θ бесконечен в обе стороны, то априорное распределение берется по-прежнему пропорциональным $d\theta$; (2) если θ изменяется от 0 до ∞, то априорное распределение принимается пропорциональным $d\theta/\theta$.

Пример 21.3

В случае нормального распределения, рассмотренного в **21.2**, статистика $\bar{x}$ достаточна для μ, и мы имеем

$$\mathbf{P}(\bar{x} \mid \mu,\ H) = \frac{1}{(2\pi)^{1/2}} \exp\left\{-\frac{n}{2}(\bar{x}-\mu)^2\right\}. \qquad (21.81)$$

Если μ лежит в интервале $(-\infty, +\infty)$, то в качестве априорного распределения берем

$$\mathbf{P}(d\mu \mid H) = d\mu. \qquad (21.82)$$

Следовательно, для апостериорного распределения параметра μ получаем

$$\mathbf{P}(d\mu \mid \bar{x},\ H) \propto \frac{1}{(2\pi)^{1/2}} \exp\left\{-\frac{n}{2}(\bar{x}-\mu)^2\right\} d\mu. \qquad (21.83)$$

Интегрируя по μ от $-\infty$ до $+\infty$, легко показать, что пропорциональность есть на самом деле равенство. Таким образом, мы можем при любом заданном уровне вероятности определить интервал для μ. Он получается тем же самым, что и в доверительно-интервальной или фидуциальной теориях.

С другой стороны, для распределения (21.6) из примера 21.2 возьмем априорное распределение параметра θ, лежащего в интервале $(0, \infty)$, в виде

$$\mathbf{P}(d\theta \mid H) = d\theta/\theta. \tag{21.84}$$

Здесь будет видно существенное сходство с фидуциальной процедурой примера 21.2. Имеем, как и там,

$$\mathbf{P}(d\theta \mid t, H) \propto \left(\frac{\beta}{\theta}\right)^{\beta} \frac{t^{\beta-1} e^{-\beta t/\theta}}{\Gamma(\beta)} \frac{d\theta}{\theta}. \tag{21.85}$$

Вычисление константы из условия, чтобы интеграл от $\theta = 0$ до ∞ равнялся единице, дает

$$\mathbf{P}(d\theta \mid t, H) = \left(\frac{\beta}{\theta}\right)^{\beta} \frac{t^{\beta} e^{-\beta t/\theta}}{\theta\Gamma(\beta)} d\theta, \tag{21.86}$$

что совпадает с распределением, полученным при доверительно-интервальном и фидуциальном подходах.

21.29 Рассмотрим теперь задачу нахождения границ для среднего в нормальных выборках, когда дисперсия неизвестна. Для распределения Стьюдента имеем

$$\mathbf{P}(dt \mid \mu, \sigma, H) = \frac{k\,dt}{(1 + t^2/\nu)^{(\nu+1)/2}}, \tag{21.87}$$

где k — некоторая постоянная, а $\nu = n - 1$. Параметры μ и σ не входят в правую часть, и, следовательно, они несущественны для $\mathbf{P}(dt|H)$ и могут быть опущены. Таким образом,

$$\mathbf{P}(dt \mid H) = \frac{k\,dt}{(1 + t^2/\nu)^{(\nu+1)/2}}. \tag{21.88}$$

Пусть теперь мы *принимаем*, что

$$\mathbf{P}(dt \mid \bar{x}, s, H) = f(t)\,dt. \tag{21.89}$$

Как и выше, $\bar{x}$ и s могут быть опущены; поэтому

$$\mathbf{P}(dt \mid H) = f(t)\,dt. \tag{21.90}$$

Сравнивая последнее соотношение с (21.88), получаем

$$\mathbf{P}(dt \mid \bar{x}, s, H) = \frac{k\,dt}{(1 + t^2/\nu)^{(\nu+1)/2}}. \tag{21.91}$$

Теперь обычным путем можно найти границы для t при заданных $\bar{x}$ и s. Джеффрейс, однако, подчеркивает, что этот результат опирается на новый постулат, выражаемый равенством

(21.89), который хотя и является естественным, вовсе не тривиален. Этот постулат равносилен следующему предположению: если мы сравниваем различные распределения, выборки из которых дают разные значения $\bar{x}$ и s, то масштаб распределения параметра μ должен быть взят пропорциональным s, а его среднее заменено разностью выборочных средних.

21.30 Аналогичным образом мы найдем, что для того, чтобы прийти к распределению Фишера — Беренса, необходимо постулировать, что

$$\mathbf{P}\{dt_1, dt_2 \mid \bar{x}_1, \bar{x}_2, s_1, s_2, H\} = f_1(t_1) f_2(t_2)\, dt_1\, dt_2. \qquad (21.92)$$

Вывод Джеффрейса распределения Фишера — Беренса из теоремы Байеса проходил бы следующим образом.

Априорное распределение $d\mu_1 d\mu_2 d\sigma_1 d\sigma_2 \mid H$ есть

$$\mathbf{P}\{d\mu_1\, d\mu_2\, d\sigma_1\, d\sigma_2 \mid H\} \propto \frac{d\mu_1\, d\mu_2\, d\sigma_1\, d\sigma_2}{\sigma_1 \sigma_2}.$$

Функция правдоподобия (наблюдения обозначены через D) имеет вид

$$\mathbf{P}\{D \mid \mu_1, \mu_2, \sigma_1, \sigma_2, H\} \propto \frac{1}{\sigma_1^{n_1} \sigma_2^{n_2}} \exp\left[-\frac{n_1}{2\sigma_1^2}\{(\mu_1 - \bar{x}_1)^2 + s_1^2\} - \right.$$
$$\left. - \frac{n_2}{2\sigma_2^2}\{(\mu_2 - \bar{x}_2)^2 + s_2^2\} \right].$$

Следовательно, по теореме Байеса

$$\mathbf{P}\{d\mu_1\, d\mu_2\, d\sigma_1\, d\sigma_2 \mid D, H\} =$$
$$= \frac{1}{\sigma_1^{n_1+1} \sigma_2^{n_2+1}} \exp\left[-\frac{n_1}{2\sigma_1^2}\{(\mu_1 - \bar{x}_1)^2 + s_1^2\} - \frac{n_2}{2\sigma_2^2}\{(\mu_2 - \bar{x}_2)^2 + s_2^2\} \right].$$

Интегрируя по σ_1 и σ_2, находим для параметров μ_1, μ_2 апостериорное распределение, которое легко привести к виду (21.75).

Обсуждение

21.31 Было так много полемики вокруг различных методов оценивания, описанных выше, что в этом месте мы оставляем обычную объективную точку зрения и сами вступаем в дискуссию. В оставшейся части этой главы будут изложены личные взгляды авторов. Мы думаем, что наша точка зрения правильна; она представляет собой результат многолетнего размышления над запутанным спорным вопросом и является серьезной попыткой понять, что говорят сторонники этих методов, и даже более серьезной попыткой догадаться, что они имеют в виду. Заслужит ли это их одобрение — мы не знаем.

21.32 Нам предстоит, таким образом, рассмотреть три подхода: доверительные, фидуциальные и байесовские интервалы.

Нас не должно вводить в заблуждение сходство результатов, получаемых в простых случаях различными методами, хотя в этом и состоит некоторое утешение. Однако мы будем придерживаться тезиса, что там, где они не совпадают, основная причина различия не в том, что тот или другой подход неверен, а в том, что они, сознательно или несознательно, либо отвечают на различные вопросы, либо основываются на разных постулатах.

21.33 Для простоты начнем с подхода Байеса — Джеффрейса. Если допустить, что вероятность является мерой доверия или некоторым неопределяемым понятием, подчиняющимся обычным постулатам, то теорема Байеса не вызывает возражений. Однако необходимо признать, что, оставляя попытку основывать теорию вероятностей на частотах событий, мы несколько теряем в объективности.

Вторая трудность состоит в принятии правил, касающихся априорных распределений вероятностей. Джеффрейс очень убедительно отстаивал правила, приведенные выше, и ничего лучшего предложено не было. Тем не менее кажется несколько произвольным, например, требование, чтобы априорное распределение параметра, изменяющегося от $-\infty$ до $+\infty$, было $d\mu$, а для параметра, лежащего в интервале от 0 до ∞, пропорционально $d\mu/\mu$. Остается впечатление, что изощренные рассуждения относительно этого различия не затрагивают корня проблемы.

21.34 Заметим также, что мы использовали байесовский подход в случаях, где существуют достаточные статистики. Но это несущественно. Если L — функция правдоподобия, то всегда можно записать

$$\mathsf{P}\{\theta \mid x_1, \ldots, x_n, H\} \propto \mathsf{P}(\theta \mid H)\, L(x_1, \ldots, x_n \mid \theta, H) \quad (21.93)$$

и, зная $\mathsf{P}(\theta|H)$, определить апостериорное распределение параметра θ. С этой точки зрения единственное преимущество малого числа достаточных статистик состоит в том, что они суммируют всю необходимую информацию, содержащуюся в функции правдоподобия, в статистики, количество которых меньше, чем n выборочных значений. Как было замечено ранее, сами выборочные значения всегда образуют множество достаточных статистик, хотя на деле это может быть всего лишь утешительная тавтология.

21.35 Доверительно-интервальная теория также является общей в том смысле, что существование единственной достаточной статистики для неизвестного параметра есть удобство, а не необходимость. Однако было уже замечено, что там, где не существует одномерной достаточной статистики, могут возникнуть мнимые или в каком-нибудь другом смысле бессодержательные

интервалы. В то же время мы не знаем случаев, когда эти трудности возникали бы при наличии одномерной достаточной статистики, так что доверительно-интервальная теория, возможно, не так свободна от потребности в достаточности, как это может показаться; но может быть лучше было бы сказать, что там, где не удается получить вложенных и просто связанных между собой интервалов, имеются специальные трудности интерпретации.

Принципиальный довод в пользу доверительных интервалов, однако, состоит в том, что они могут быть выведены в терминах частотной теории вероятностей без каких-либо предположений, касающихся априорного распределения, которые так существенны для байесовского подхода. Это, на наш взгляд, неоспоримо. Но справедливо будет спросить, достигают ли они указанной экономии основных предположений без потери чего-либо, свойственного байесовской теории. Наше мнение таково, что кое-что действительно иногда теряется и это кое-что может быть важным для задач оценивания.

21.36 Рассмотрим пример, где оценивается среднее μ нормальной совокупности при известной дисперсии. Пусть нам *известно*, что μ лежит между 0 и 1. Согласно постулату Байеса имеем

$$\mathbf{P}(d\mu \mid \bar{x}) = \frac{\exp\left\{-\frac{n}{2}(\mu - \bar{x})^2\right\} d\mu}{\int_0^1 \exp\left\{-\frac{n}{2}(\mu - \bar{x})^2\right\}^2 d\mu}, \qquad (21.94)$$

и задача определения границ для μ допускает решение, хотя и остались математические трудности. Что в этом случае способна дать доверительно-интервальная теория? Она может лишь повторить утверждения вида

$$\mathbf{P}\{\bar{x} - 1{,}96/\sqrt{n} \leqslant \mu \leqslant \bar{x} + 1{,}96/\sqrt{n}\} = 0{,}95.$$

Они по-прежнему справедливы в требуемой части случаев, но эти утверждения не учитывают наших априорных знаний об интервале изменения μ и иногда оказываются бесполезными. К примеру, возможно, верное утверждение $-1 \leqslant \mu \leqslant 2$ будет абсурдным, если мы уже знаем, что $0 \leqslant \mu \leqslant 1$. Конечно, полученный интервал можно урезать в соответствии с априорными данными. В приведенном примере остается только утверждать, что $0 \leqslant \mu \leqslant 1$, и наблюдения ничего не прибавляют к нашим знаниям.

В действительности, как нам кажется, доверительно-интервальная теория имеет дефект в ее главном достоинстве: она достигает общности ценой того, что оказывается неспособной

включать априорные знания в свои утверждения. Когда мы делаем окончательный вывод о μ, то должны синтезировать информацию, полученную из наблюдений, с нашими априорными сведениями. Теорема Байеса имеет целью достичь указанного синтеза с самого начала. Доверительная теория оставляет это на самый конец (а ее наиболее распространенные изложения игнорируют этот момент полностью).

21.37 Фидуциальная теория, как мы уже заметили, была предложена Фишером для случая, когда имеются достаточные статистики, или, в общем виде, когда может быть использована вся информация, входящая в функцию правдоподобия. Никакого систематического изложения процедур, которым нужно следовать в задачах, где доступна априорная информация, не давалось. Но, по-видимому, нет причин, чтобы нельзя было использовать метод, подобный тому, что проиллюстрирован уравнением (21.94). Иначе говоря, если мы получили фидуциальное распределение $f(\mu)$, сосредоточенное на произвольном интервале, и имеется дополнительная информация о том, что параметр лежит в интервале от μ_0 до μ_1 (внутри исходного интервала), то с помощью усечения мы модифицируем фидуциальное распределение к виду

$$f(\mu) \Big/ \int_{\mu_0}^{\mu_1} f(\mu)\, d\mu.$$

21.38 Одну из критикуемых трудностей фидуциальной теории иллюстрирует вывод распределения Стьюдента в фидуциальной форме, приведенный в **21.10**. Нам кажется, что этот вопрос был неверно понят всеми, кроме Джеффрейса. Поскольку распределение Стьюдента дает одинаковый результат как для фидуциальной, так и для доверительной теорий, тогда как различаются эти два подхода в задаче о двух средних, то обе стороны, кажется, искали свои основные различия во второй задаче, а не в первой. На наш же взгляд, *c'est le premier test qui coûte* *). Если принять ход рассуждений, применяемый в первой задаче, то более общий результат Фишера — Беренса получается очень простым обобщением. Это также понятно из подхода Байеса — Джеффрейса, при котором соотношение (21.92) является очевидным обобщением (21.90) на случай двух независимых выборок.

Как отмечалось в **21.10**, вопрос состоит в следующем. Можем ли мы в совместном распределении величин $\bar{x}$ и s (которые независимы в обычном смысле) заменить $d\bar{x}\, ds$ на $d\mu\, d\sigma/\sigma$? Нам кажется, что последнее не очевидно и что для этого на самом

*) Именно первый критерий является ценным (фр.). (*Прим. перев.*)

деле, точно так же как и для соотношения (21.90), требуется новый постулат. Этот вопрос был отчетливо поставлен в работе Иэйтса (1939а). Глубокое общее обсуждение фидуциальной теории дано Демпстером (1964) и в работах Симпозиума, опубликованных в Бюллетене Международного статистического института (1964, **40(2)**, стр. 833—939).

Парадоксы и ограничения фидуциальной теории

21.39 Если (t_1, t_2) совместно достаточны для (θ_1, θ_2), то можно написать альтернативные факторизации

$$L(x \mid \theta_1, \theta_2) \propto g(t_1, t_2 \mid \theta_1, \theta_2) = g_1(t_1 \mid t_2, \theta_1, \theta_2)\, g_2(t_2 \mid \theta_1, \theta_2) = \\ = g_3(t_2 \mid t_1, \theta_1, \theta_2)\, g_4(t_1 \mid \theta_1, \theta_2).$$

Если каждая из статистик t_1 и t_2 зависит только от одного из параметров, то никаких трудностей не возникает, поскольку каждая из них является одномерной достаточной статистикой для своего параметра.

В более общем случае мы можем различать две специальные структуры достаточных статистик.

(а) Одна из статистик зависит только от одного параметра. Тогда факторизация принимает вид

$$\left.\begin{aligned} &L \propto g_1(t_1 \mid t_2, \theta_1, \theta_2)\, g_2(t_2 \mid \theta_2), \\ \text{либо} \quad &L \propto g_3(t_2 \mid t_1, \theta_1, \theta_2)\, g_4(t_1 \mid \theta_1). \end{aligned}\right\} \qquad (21.95)$$

(б) Одно из условных распределений зависит только от одного параметра. Имеем

$$\left.\begin{aligned} &L \propto g_1(t_1 \mid t_2, \theta_1)\, g_2(t_2 \mid \theta_1, \theta_2) \\ \text{либо} \quad &L \propto g_3(t_2 \mid t_1, \theta_2)\, g_4(t_1 \mid \theta_1, \theta_2). \end{aligned}\right\} \qquad (21.96)$$

Если выполняется первое из соотношений (21.96), то t_2 является одномерной достаточной статистикой для θ_2 при известном θ_1; если же имеет место второе соотношение, то t_1 есть одномерная достаточная статистика для θ_1, когда известно θ_2.

Любая строка в соотношениях (21.95) или (21.96) позволяет строить совместное фидуциальное распределение следующим образом. Вначале находим фидуциальное распределение одного параметра из множителя, в который входит только этот параметр, и затем получаем условное фидуциальное распределение другого параметра (при фиксированном значении первого) из множителя, содержащего оба параметра. Произведение этих распределений принимается за совместное фидуциальное распределение, по аналогии с теоремой об умножении вероятностей. Именно так Фишер (1956) и Кэнуй (1958) использовали соотношения (21.95) и (21.96) соответственно.

Мы видели, что в **21.10** и **21.26** соотношения (21.95) и (21.96) выполнялись. Это было возможно потому, что выборочное среднее (или разность средних) t_1 не зависело от выборочной дисперсии (или дисперсий) t_2. В общем случае, однако, даже эти специальные структуры достаточных статистик недостаточны для того, чтобы гарантировать единственность совместного фидуциального распределения, как это показали контрпримерами Тьюки (1957) и Бриллинджер (1962). Неединственность возникает в точности тогда, когда имеют место одновременно *оба* соотношения в (21.95) (или в (21.96)), и, следовательно, совместное фидуциальное распределение зависит от того, какое из них мы используем для его построения. См. также Молдон (1955) и Демпстер (1963).

В **23.37—39** мы увидим, что при определенных условиях каждая из структур достаточности (21.95), (21.96) обеспечивает оптимальные свойства для условных критериев.

Фрэзер (1961a, b) обсуждал связь между фидуциальными выводами и некоторыми свойствами инвариантности. См. также Хора и Бюлер (1966).

21.40 Линдли (1958a) получил простой, но имеющий большие последствия результат, который не только проливает свет на связь между фидуциальным и байесовским подходами, но и ограничивает претензии фидуциальной теории на то, что она дает общий метод, не противоречащий байесовскому и сочетающийся с ним. В действительности, как показал Линдли, фидуциальный подход не противоречит байесовским методам тогда и только тогда, когда он применяется к случайной величине x и параметру θ, которые могут быть (порознь) преобразованы в u и τ соответственно, так что τ является параметром расположения для u. В этом случае фидуциальный подход эквивалентен байесовскому с равномерным априорным распределением параметра τ. Приведенная критика относится в равной мере и к «доверительным распределениям», определенным выше, в конце **20.6**, поскольку они совпадают с фидуциальными распределениями.

21.41 Используя равенство (21.5), запишем фидуциальное распределение параметра θ в виде (не путать с обозначением характеристической функции)

$$\varphi_x(\theta) = -\frac{\partial}{\partial\theta} F(x \mid \theta), \tag{21.97}$$

в то время как апостериорное распределение величины θ, при заданном априорном распределении $p(\theta)$, по теореме Байеса есть

$$\pi_c(\theta) = p(\theta) f(x \mid \theta) \Big/ \int p(\theta) f(x \mid \theta)\, d\theta, \tag{21.98}$$

где $f(x|\theta) = \partial F(x|\theta)/\partial x$ — функция плотности. Обозначим $r(x)$ знаменатель в правой части (21.98). Тогда

$$\pi_x(\theta) = \frac{p(\theta)}{r(x)} \frac{\partial F(x|\theta)}{\partial x}. \tag{21.99}$$

Если имеется некоторое априорное распределение $p(\theta)$, для которого фидуциальное распределение эквивалентно байесовскому апостериорному распределению, то (21.97) и (21.99) будут равны, т. е.

$$\frac{-\frac{\partial}{\partial\theta} F(x|\theta)}{\frac{\partial}{\partial x} F(x|\theta)} = \frac{p(\theta)}{r(x)}. \tag{21.100}$$

Выражение (21.100) показывает, что отношение в его левой части должно быть произведением функции от θ на функцию от x. Перепишем это так:

$$\frac{1}{r(x)} \frac{\partial F}{\partial x} + \frac{1}{p(\theta)} \frac{\partial F}{\partial \theta} = 0. \tag{21.101}$$

При заданных $p(\theta)$ и $r(x)$ определим F из уравнения (21.101). Единственное решение, отличное от постоянной, имеет вид

$$F = G\{R(x) - P(\theta)\}, \tag{21.102}$$

где G — произвольная функция, а R и P являются соответственно интегралами от r и p относительно их аргументов. Если обозначим $u = R(x)$, $\tau = P(\theta)$, то (21.102) преобразуется к виду

$$F = G\{u - \tau\}, \tag{21.103}$$

так что τ является параметром расположения для u. Обратно, если выполняется соотношение (21.103), то выполнено (21.100) с u и τ вместо x и θ и с равномерным распределением $p(\tau)$. Таким образом, (21.103) есть необходимое и достаточное условие для того, чтобы имело место равенство (21.100), т. е. для того, чтобы фидуциальное распределение было эквивалентно некоторому байесовскому апостериорному распределению.

21.42 Рассмотрим теперь случай, когда имеются две независимые выборки, информация о которых заключена в достаточных статистиках x и y, на основе которых нужно сделать вывод о параметре θ. Для этого существует два пути:

(а) мы можем рассмотреть обе выборки одновременно и получить фидуциальное распределение $\varphi_{x,y}(\theta)$;

(б) можно получить фидуциальное распределение $\varphi_x(\theta)$ из первой выборки и, используя его как априорное распределение при байесовском подходе, на основе второй выборки найти апостериорное распределение $\pi_{x,y}(\theta)$.

Если фидуциальный подход не противоречит байесовскому, то (а) и (б) логически эквивалентны, и мы должны получить $\varphi_{x,\,y}(\theta) = \pi_{x,\,y}(\theta)$.

Возьмем простейший случай, когда x и y одинаково распределены. Поскольку их распределение допускает одномерную достаточную статистику для θ, то исходная функция плотности имеет вид (17.83). Поэтому можно считать (см. упражнение 17.14), что распределение величины x само представимо в виде

$$f(x \mid \theta) = f(x)\, g(\theta) \exp(x\theta), \tag{21.104}$$

и аналогично для y во второй выборке. Более того, в объединенной выборке величина $x + y$ является, очевидно, достаточной статистикой для θ, и тогда комбинированное фидуциальное распределение $\varphi_{x,y}(\theta)$ есть функция только от $x + y$ и θ. Найдем теперь условия, при которых $\pi_{x,y}(\theta)$ также будет функцией только от $x + y$ и θ. По теореме Байеса

$$\pi_{x,\,y}(\theta) = -\frac{\varphi_x(\theta)\, f(y \mid \theta)}{\int \varphi_x(\theta)\, f(y \mid \theta)\, d\theta}$$

и если $\pi_{x,\,y}(\theta)$ есть функция только от $x + y$ и θ, то тем же свойством обладает и отношение при двух различных значениях параметра θ:

$$\frac{\pi_{x,\,y}(\theta)}{\pi_{x,\,y}(\theta')} = \frac{\varphi_x(\theta)\, f(y \mid \theta)}{\varphi_x(\theta')\, f(y \mid \theta')}. \tag{21.105}$$

Таким образом, соотношение (21.105) должно быть инвариантным относительно перестановки x и y. Следовательно, подставляя (21.104) в (21.105), получаем

$$\frac{\varphi_x(\theta)\, g(\theta)}{\varphi_x(\theta')\, g(\theta')} \exp\{y(\theta - \theta')\} = \frac{\varphi_y(\theta)\, g(\theta)}{\varphi_y(\theta')\, g(\theta')} \exp\{x(\theta - \theta')\},$$

так что

$$\frac{\varphi_x(\theta)}{\varphi_x(\theta')} \frac{\varphi_y(\theta')}{\varphi_y(\theta)} = \exp\{(x - y)(\theta - \theta')\},$$

или

$$\varphi_x(\theta) = \frac{\varphi_x(\theta')\, e^{-x\theta'}\, \varphi_y(\theta)\, e^{-y\theta}}{\varphi_y(\theta')\, e^{-y\theta'}}\, e^{x\theta}. \tag{21.106}$$

Если считать θ' и y постоянными, то (21.106) можно переписать в виде

$$\varphi_x(\theta) = A(x)\, B(\theta)\, e^{x\theta}, \tag{21.107}$$

где A и B — произвольные функции. Из (21.97), (21.104) и (21.107) имеем

$$\frac{-\frac{\partial}{\partial\theta}F(x\mid\theta)}{\frac{\partial}{\partial x}F(x\mid\theta)}=\frac{\varphi_x(\theta)}{f(x\mid\theta)}=\frac{A(x)\,B(\theta)}{f(x)\,g(\theta)}. \tag{21.108}$$

Но (21.108) совпадает в точности с соотношением (21.100), для которого (21.103) является необходимым и достаточным условием. Таким образом, $\varphi_{x,y}(\theta)=\pi_{x,y}(\theta)$ возможно тогда и только тогда, когда x и θ можно преобразовать в (21.103), причем τ будет параметром расположения для u, а $p(\tau)$ — равномерным распределением. Итак, фидуциальный подход совпадает с байесовским тогда и только тогда, когда задачу можно свести к задаче о параметре расположения с равномерным априорным распределением. Пример, где это не так, дан в упражнении 21.11.

Линдли показал также, что в экспоненциальном семействе распределений (17.83) только нормальное и гамма-распределение удовлетворяют условию о возможности преобразования к виду (21.103): этим объясняется идентичность результатов, полученных фидуциальным и байесовским методами для указанных случаев (см. пример 21.3). Спротт (1960, 1961) показал, что этот результат остается в силе, когда x и y неодинаково распределены.

Уэлч и Пирс (1963), Уэлч (1965) и Пирс (1965) изучали соответствие между байесовскими и доверительными интервалами, главным образом для асимптотических решений. Тэтчер (1964) рассмотрел это соответствие для биномиальных прогнозов. Гейссер и Корнфилд (1963) и Фрэзер (1964) указали на дальнейшие трудности с фидуциальными распределениями в многомерном случае. Смотри также Работы Симпозиума Международного статистического института, упоминавшиеся в конце **21.38**.

Фрэзер (1962) предложил модификацию фидуциального метода, которая расширяет область его непротиворечивости с байесовскими методами.

21.43 Имеется еще одно возражение против фидуциальной теории, которое уже упоминалось в связи с байесовским подходом: она отказывается от строго частотного подхода в задаче интервального оценивания. На самом деле возможно, как показал Барнард (1950), обосновать решение Фишера — Беренса задачи о двух средних с некоторой иной частотной точки зрения, но, как он сам утверждает далее, идея фиксированного пространства элементарных исходов, в терминах которого интерпретируются частоты, чужда фидуциальному подходу. Таким образом, статистику предоставляется выбор между доверитель-

ными интервалами, дающими точные интерпретируемые в частотах утверждения, которые в исключительных случаях могут оказаться тривиальными, и другими методами, отказывающимися от частотной интерпретации в пользу выводов, которые представляются, возможно интуитивно, более относящимися к делу.

УПРАЖНЕНИЯ

21.1 Пусть $\bar{x}$ — среднее для выборки объема n из совокупности с распределением

$$dF = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx,$$

пусть $s'^2 = \sum (x-\bar{x})^2/(n-1)$, а x — еще одно независимое выборочное значение. Доказать, что

$$t = \frac{x-\bar{x}}{s'}\sqrt{\frac{n}{n+1}}$$

имеет распределение Стьюдента с $n-1$ ст. св. Показать далее, что фидуциальные границы для x есть

$$\bar{x} \pm s' t_1 \sqrt{\frac{n+1}{n}},$$

где t_1 выбрано так, чтобы интеграл от распределения Стьюдента в интервале от $-t_1$ до t_1 был равен заданному значению $1-\alpha$.

(Фишер, 1935b. Этот результат дает оценку для следующего значения, когда уже имеется n наблюдений, и распространяет идею фидуциальных границ с параметров на случайные величины, зависящие от них.)

21.2 Аналогично предыдущему показать, что если выборка объема n_1 дает среднее $\bar{x}_1$ и оценку дисперсии $s_1'^2$, то фидуциальное распределение среднего $\bar{x}_2$ и оценки дисперсии $s_2'^2$, полученных из второй выборки объема n_2, есть

$$dF \propto \frac{s_1'^{(n_1-1)}\, s_2'^{(n_2-1)}\, d\bar{x}_2\, ds_2'}{\left\{(n_1-1)s_1'^2 + (n_2-1)s_2'^2 + (\bar{x}_1-\bar{x}_2)^2 \frac{n_1 n_2}{n_1+n_2}\right\}^{(n_1+n_2-1)/2}}.$$

Вывести отсюда, устремляя n_2 к бесконечности, совместное фидуциальное распределение для μ и σ.

(Фишер, 1935b.)

21.3 Пусть кумулятивная биномиальная функция распределения дается равенством

$$G(f, \pi) = \sum_{i=f}^{n} \binom{n}{j} \pi^j (1-\pi)^{n-i}.$$

Доказать, что f/n есть достаточная статистика для π и

$$h_0(\pi)\,d\pi \equiv \frac{\partial G(f,\pi)}{\partial \pi}\,d\pi = \binom{n}{f-1}\pi^{f-1}(1-\pi)^{n-f}\,d\pi$$

является допустимым фидуциальным распределением для π. Показать, что

$$h_1(\pi)\,d\pi \equiv \frac{\partial G(f+1,\pi)}{\partial \pi}\,d\pi = \binom{n}{f}\pi^{f}(1-\pi)^{n-f-1}\,d\pi$$

есть тоже допустимое распределение. Показать, как определить π_0 из h_0 и π_1 из h_1, чтобы фидуциальный интервал $\pi_0 \leqslant \pi \leqslant \pi_1$ имел вероятность, не меньшую заданного $1-\alpha$.

(Стивенс, 1950. Использование двух фидуциальных распределений необходимо из-за дискретности наблюдений величины f. Ср. **20.22** по поводу аналогичной трудности в доверительных интервалах.)

21.4 Пусть $l_{11}, l_{12}, \ldots, l_{1,n-1}$ представляют собой $n-1$ линейных функций от наблюдений, ортогональных к $\bar{x}_1$ и между собой, имеющих среднее нуль и дисперсию σ_1^2. Аналогично определим $l_{21}, l_{22}, \ldots, l_{2,n-1}$.

Тогда при двух выборках объема n из нормальных популяций с одинаковыми средними и дисперсиями σ_1^2, σ_2^2 функция

$$\frac{(\bar{x}_1-\bar{x}_2)\,n^{1/2}}{\left\{\sum (l_{1j}+l_{2j})^2/(n-1)\right\}^{1/2}}$$

будет иметь распределение Стьюдента t с $n-1$ ст. св. Показать, как, используя этот результат, строить доверительные интервалы для разности двух средних. Показать также, что (21.51) при $n_1 = n_2$ является статистикой такого же вида.

21.5 Пусть даны две выборки объемов n_1 и n_2 из нормальных совокупностей с неравными дисперсиями. Показать, что, выбирая случайным образом n_1 элементов из n_2 (где $n_1 \leqslant n_2$) и случайным образом объединяя их в пары с элементами первой выборки, можно получить доверительные интервалы для разности между средними на основе распределения Стьюдента независимо от отношения дисперсий популяций. Показать, что это эквивалентно тому, чтобы положить $c_{ij} = 0$ $(i \neq j)$; $=1$ $(i = j)$ в (21.36), и, следовательно, это неэффективное решение задачи о двух средних.

21.6 Используя метод из **21.23**, показать, что статистика (21.21) имеет приближенно распределение Стьюдента с числом степеней свободы, определяемым соотношением (21.64).

21.7 Исходя из F-распределения Фишера (16.24), найти фидуциальное распределение для $\theta = \sigma_1^2/\sigma_2^2$. Показать, что если рассматривать распределение Стьюдента статистики (21.20) как совместное фидуциальное распределение для δ и θ и проинтегрировать его по фидуциальному распределению θ, то мы получим результат пункта **21.26** о распределении статистики z.

(Фишер, 1939.)

21.8 Доказать утверждение из **21.16** о том, что если $ax + by = z$, где x и y — независимые случайные величины и каждая из величин x, y, z имеет χ^2-распределение, то $a = b = 1$.

(Шеффе, 1944.)

21.9 Показать, что если взять первые два члена в правой части разложения (21.71), то соотношение (21.65) будет, с точностью до $1/n$, аппроксимацией, приведенной в **21.24**, для распределения статистики (21.21), т. е. будет распределением Стьюдента с числом степеней свободы, задаваемым (21.64).

21.10 Доказать, что для $n_1 = n_2 = n$ условное распределение статистики z из (21.21) при фиксированном отношении s_1/s_2 можно получить, используя тот факт, что величина

$$\left(\frac{2}{\dfrac{1-\left(s_1^2/s_2^2\right)}{1-\left(\sigma_1^2/\sigma_2^2\right)}+\dfrac{s_1^2/s_2^2}{\sigma_1^2/\sigma_2^2}}\right)^{1/2} z$$

имеет t-распределение Стьюдента с $2(n-1)$ ст. св.

(Бартлетт, 1936)

21.11 Показать, что если достаточная статистика x имеет распределение

$$f(x\,|\,\theta)=\frac{\theta^2}{\theta+1}(x+1)\,e^{-x\theta}, \qquad x>0, \quad \theta\geqslant 0,$$

то фидуциальное распределение параметра θ для объединения выборок с достаточными статистиками x, y имеет вид

$$\varphi_{x,\,y}(\theta)=\frac{e^{-z\theta}}{(\theta+1)^3}\left[\theta^3\left(2z^2+\frac{4}{3}z^3+\frac{1}{6}z^4\right)+\theta^4\left(z^2+z^3+\frac{1}{6}z^4\right)\right]$$

(где $z=x+y$), тогда как для одной выборки

$$\varphi_x(\theta)=\frac{\theta x e^{-x\theta}}{(\theta+1)^2}\,[1+(1+\theta)(1+x)].$$

(Заметим, что знак минус в соотношении (21.5) не является здесь необходимым, поскольку $F(x|\theta)$ по θ является возрастающей функцией.) Показать далее, что байесовское апостериорное распределение, полученное по второй выборке при использовании $\varphi_x(\theta)$ как априорного распределения, представимо в виде

$$\pi_{x,\,y}(\theta)\propto e^{-z\theta}\left(\frac{\theta}{\theta+1}\right)^3 x(1+y)\,[1+(1+\theta)(1+x)],$$

так что $\pi_{x,\,y}(\theta)\not\equiv\varphi_{x,\,y}(\theta)$. Заметим также, что $\pi_{x,\,y}(\theta)\not\equiv\pi_{y,\,x}(\theta)$.

(Линдли, 1958а.)

ГЛАВА 22

ПРОВЕРКА ГИПОТЕЗ: ПРОСТЫЕ ГИПОТЕЗЫ

22.1 Теперь мы переходим от задач оценивания параметров к задачам проверки гипотез о параметрах. Наша цель будет заключаться уже не в том, чтобы найти наилучшую (точечную или интервальную) оценку неизвестного параметра, а в том, чтобы решить, приемлема ли, в свете полученных наблюдений, некоторая известная величина в качестве значения параметра.

В некотором смысле задача проверки гипотез логически предшествует задаче оценивания. Например, при исследовании разности между средними двух нормальных совокупностей нам в первую очередь нужно решить, указывают ли наблюдения на наличие *какой бы то ни было* отличной от нуля разности между средними. Другими словами, мы должны сравнить наблюденные различия между выборками с тем, чего можно было бы ожидать при гипотезе, что средние совокупностей совпадают и что все различия объясняются выборочными флуктуациями. Если эта гипотеза не подтвердится, тогда надо переходить к следующему этапу, имеющему целью оценивание *величины* разности между средними.

Ясно, что задачи проверки гипотез тесно связаны с задачами оценивания. Однако полезно рассмотреть их отдельно, по крайней мере ради упрощения изложения. Многие из идей, изложенных в этой и следующей главе, принадлежат Нейману и Э. Пирсону, авторам фундаментальной серии работ, посвященных теории проверки гипотез (1928, 1933a, b, 1936a, 1938). См. также монографию Лемана (1959).

22.2 Понятие гипотезы, с которым имеет дело статистика, более узко, чем общее понятие научной гипотезы. Так, существует научная гипотеза, что любая материальная частица во Вселенной притягивает любую другую или что существует жизнь на Марсе. Однако эти гипотезы не относятся к числу тех, которые проверяются в статистике. Статистические гипотезы касаются поведения наблюдаемых случайных величин. Более точно, пусть имеется набор случайных величин $x_1, \ldots, x_n$. Этот набор можно представить в виде точки $\boldsymbol{x}$ в n-мерном выборочном пространстве, каждой из координатных осей которого сопоставлена одна

случайная величина. Случайной величине x соответствует некоторое распределение вероятностей, и для любой области w в выборочном пространстве W можно (по крайней мере в принципе) вычислить вероятность $\mathsf{P}(x \in w)$ того, что выборочная точка x попадет в w. Любая гипотеза, связанная с $\mathsf{P}(x \in w)$, будет называться статистической.

Примерами статистических гипотез могут служить следующие гипотезы: (а) нормальное распределение имеет заданные среднее и дисперсию; (б) нормальное распределение имеет заданное среднее (о дисперсии ничего не говорится); (в) распределение нормально (с какими-нибудь средним и дисперсией); (г) два неизвестных непрерывных распределения одинаковы. Каждой из приведенных четырех гипотез соответствуют определенные свойства выборочного пространства. Поэтому каждая из них может быть сформулирована в виде утверждения, относящегося к выборочному пространству, и проверена сравнением с наблюдениями.

Параметрические и непараметрические гипотезы

22.3 Заметим, что в примерах (а) и (б) последнего пункта распределение, порождающее наблюдения, имело известный вид (было нормальным) и гипотезы касались только значений одного или обоих параметров распределения. Такие гипотезы называются *параметрическими.*

Гипотеза (в) имеет другую природу. Она эквивалентна гипотезе о том, что все семиинварианты распределения конечны и что семиинварианты порядка выше второго равны нулю (см. пример 3.11). Термин «параметр» часто используется для обозначения семиинварианта или момента распределения, чтобы отличить их от соответствующих выборочных величин. Такое употребление понятно, однако оно неточно. Нормальное распределение

$$dF(x) = (2\pi)^{-1/2} \exp\left\{ -\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \right\} \frac{dx}{\sigma}$$

имеет два параметра, μ и σ. (Иногда удобнее считать параметрами μ и σ^2 — это зависит от соглашения. Несущественные изменения такого рода не меняют *числа* параметров.) Известно, что среднее нормального распределения равно μ, а его дисперсия равна σ^2. Однако среднее и дисперсия никак не могут считаться параметрами этого распределения с бóльшим основанием, чем, скажем, медиана (также равная μ) и среднее отклонение от среднего ($=\sigma(2/\pi)^{1/2}$) или любая другая интересующая нас константа из бесконечного множества констант, включающего все моменты и семиинварианты. Под «параметрами»,

следовательно, мы понимаем конечное число констант, определяющих распределение.

При таком понимании термина «параметр» гипотезы (в) и (г) пункта **22.2** следует считать *непараметрическими*. Непараметрические гипотезы будут подробно рассмотрены в главах 30, 31, но значительная часть теории настоящей и следующей глав одинаково применима как к параметрическому, так и к непараметрическому случаям. Однако все конкретные рассмотрения будут большей частью относиться к параметрическим гипотезам.

Простые и сложные гипотезы

22.4 Есть разница между гипотезами (а) и (б) пункта **22.2**. Гипотеза (а) определяет значения всех параметров распределения, тогда как гипотеза (б) определяет только некоторое подмножество значений параметров. Для теории это различие существенно. В общем случае оно формулируется следующим образом. Если распределение имеет всего l параметров и гипотеза утверждает, что k из них имеют заданные значения, то она называется *простой*, когда $k = l$, и *сложной*, когда $k < l$. Геометрически можно представлять возможные значения параметров в виде области в пространстве размерности l (одно измерение для каждого параметра). Если гипотезой выделяется единственная точка этого параметрического пространства, то это простая гипотеза; если же гипотеза выделяет область, содержащую более одной точки, то это сложная гипотеза.

При этом $l - k$ называется *числом степеней свободы* гипотезы, а k — числом *ограничений*, налагаемых гипотезой. Эта терминология очевидным образом связана с рассмотренной геометрической иллюстрацией.

Критические области и альтернативные гипотезы

22.5 Чтобы проверить какую-либо гипотезу, исходя из (случайной) выборки наблюдений, мы должны разделить выборочное пространство (т. е. возможные наборы наблюдений) на две области. Если наблюдаемая выборочная точка $\boldsymbol{x}$ попадает в одну из этих областей, скажем w, то гипотеза отвергается; если же $\boldsymbol{x}$ попадает в дополнительную область $W - w$, то гипотеза принимается. Область w называется *критической областью* критерия, а $W - w$ называется *областью принятия*.

Следует сказать, что довольно безапелляционные термины «отвергнуть» и «принять», использованные выше, в настоящее время общеприняты, и мы также будем ими пользоваться; но они вовсе не означают, что какая бы то ни было гипотеза в науке

может быть окончательно принята или отвергнута. Читатель, которому трудно принять эти, может быть, неподходящие термины, возможно, согласится рассматривать их как условные сокращения: «отвергнуть» вместо «решить, что наблюдения говорят не в пользу» и «принять» вместо противоположного утверждения. Далее мы будем изучать процедуры, приводящие к тому или другому из этих решений с определенной вероятностью ошибки.

22.6 Если известно распределение вероятностей наблюдений, соответствующее проверяемой гипотезе H_0, то можно определить w так, чтобы при выполнении гипотезы H_0 вероятность отвергнуть гипотезу (т. е. вероятность попадания x в w) была равна заранее заданной величине α, т. е.

$$\mathbf{P}\{x \in w \mid H_0\} = \alpha. \tag{22.1}$$

В случае дискретного распределения выполнение (22.1) для любого α из интервала (0, 1) может оказаться невозможным. Значение α называется *размером* критерия *). Начиная с этого момента α будет считаться каким-то образом выбранным. Вопрос об этом выборе мы обсудим позже.

Очевидно, мы можем в общем случае найти много (часто даже бесконечно много) подобластей w выборочного пространства, для которых выполнено (22.1). Какая из них предпочтительней? Это основная проблема теории проверки гипотез. Другими словами, нужно решить, какие совокупности наблюдений говорят в пользу данной гипотезы, а какие — против.

22.7 Поставив вопрос таким образом, мы подошли к сути проблемы. Действительно, совершенно бесполезно знать только свойства критической области, соответствующие гипотезе H_0. Что произойдет, когда выполнена какая-то другая гипотеза? Иными словами, невозможно ответить на вопрос, говорят ли полученные наблюдения в пользу данной гипотезы, если мы не знаем, с какими альтернативными гипотезами она сравнивается. Вполне может случиться, что выборка довольно «неправдоподобна», если справедлива исходная гипотеза: однако она может быть еще более «неправдоподобной», если выполнена другая гипотеза. Если ситуация такова, что мы должны выбрать либо одну гипотезу, либо другую, мы, очевидно, выберем первую, несмотря на «неправдоподобность» наблюдений. Задача проверки гипотезы, в сущности, состоит в выборе между ней и

*) Проверяемая гипотеза часто называется «нулевой гипотезой», а размер критерия — «уровнем значимости». Мы не будем использовать эти термины, поскольку слова «нулевой» и «значимость» могут иногда вводить в заблуждение.

другими гипотезами. Отсюда немедленно следует, что решение о принятии или непринятии исходной гипотезы существенно зависит от того, против каких альтернативных гипотез она проверяется.

Мощность критерия

22.8 Рассуждения пункта **22.7** приводят нас к выводу, что суждение о качестве критической области (или, что то же самое, критерия) должно основываться на свойствах, которыми она обладает как в случае выполнения проверяемой гипотезы, так и в случае ее невыполнения. Из этого следует, что ошибки, возникающие при проверке статистической гипотезы, могут быть двух типов:

(I) можно ошибочно отвергнуть гипотезу, когда она верна;

(II) можно ее ошибочно принять, когда она неверна.

Эти ошибки называются соответственно ошибками первого и второго рода. Вероятность ошибки первого рода равна размеру используемой критической области. Вероятность ошибки второго рода, обозначаемая обычно β, зависит, конечно, от рассматриваемой альтернативной гипотезы H_1. Таким образом,

$$\mathsf{P}\{x \in W - w \mid H_1\} = \beta$$

или

$$\mathsf{P}\{x \in w \mid H_1\} = 1 - \beta. \qquad (22.2)$$

Вероятность $1 - \beta$ называется *мощностью* критерия для проверки гипотезы H_0 против альтернативной гипотезы H_1. Указание гипотезы H_1 в последней фразе существенно, так как мощность зависит от H_1.

Пример 22.1

Рассмотрим задачу проверки гипотезы о среднем значении нормального распределения с единичной дисперсией. Говоря точнее, требуется проверить гипотезу

$$H_0\colon\ \mu = \mu_0$$

в случае распределения

$$dF(x) = (2\pi)^{-1/2} \exp\left\{-\tfrac{1}{2}(x-\mu)^2\right\} dx, \qquad -\infty \leqslant x \leqslant \infty.$$

Это простая гипотеза, так как она определяет $F(x)$ полностью. В качестве альтернативы мы возьмем также простую гипотезу

$$H_1\colon\ \mu = \mu_1 > \mu_0.$$

Таким образом, задача заключается в том, чтобы выбрать в качестве среднего значения либо меньшее из двух заданных значений (μ_0), либо большее (μ_1).

В случае выборки объема $n = 2$ ситуацию можно представить графически. На рис. 22.1 показаны возможные совокупности выборочных точек при большом числе повторений. Нижнее скопление соответствует выполнению гипотезы H_0, а верхнее — H_1.

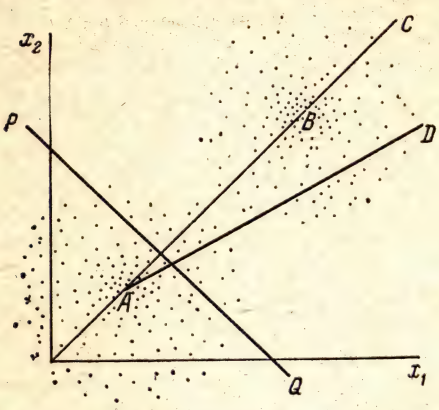


Рис. 22.1. Критические области для $n = 2$ (см. текст).

В рассматриваемом случае, конечно, выборочные распределения непрерывны, однако расположение точек дает представление о том, как выборочные плотности сгущаются около истинных средних.

Чтобы выбрать критическую область, мы должны в соответствии с (22.1) найти область на плоскости, содержащую долю α распределения H_0. Одна из таких областей представлена частью плоскости, лежащей выше прямой PQ, которая перпендикулярна к прямой AB, соединяющей гипотетические средние. (A изображает точку (μ_0, μ_0), а B — точку (μ_1, μ_1). Другой возможной критической областью размера α будет область CAD.

Из круговой симметрии скоплений сразу видно, что первая из критических областей содержит гораздо большую часть скопления H_1, чем область CAD. Поэтому при гипотезе H_1 доля случаев, когда отвергается H_0, для первой области больше, чем для второй. Следовательно, значение $1 - \beta$ для первой критической области, т. е. ее мощность, будет больше.

22.9 Пример 22.1 указывает нам очевидный критерий для выбора критической области среди всех областей, удовлетворяющих (22.1). Надо взять такую критическую область w, которая имела бы максимальную мощность (22.2). В этом случае при заданной вероятности ошибки первого рода α мы будем иметь минимальную вероятность ошибки второго рода β. Эта фундаментальная идея, впервые в явном виде сформулированная Дж. Нейманом и Э. Пирсоном, лежит в основе теории этой и следующих глав.

Критическая область, мощность которой не меньше мощности любой другой области того же размера, предназначенной

для проверки гипотезы H_0 против альтернативной гипотезы H_1, называется наилучшей критической областью (сокращенно: НКО), а критерий, основанный на НКО, называется наиболее мощным (сокращенно: НМ) критерием.

Проверка простой гипотезы H_0 против простой альтернативы H_1

22.10 Если проверяется простая гипотеза против простой альтернативной гипотезы, т. е. делается выбор между двумя полностью определенными распределениями, то задача нахождения НКО не вызывает особых трудностей. Ее решение дается леммой Неймана — Пирсона (1933b), которую мы сейчас докажем.

Как и в предыдущих главах, $L(x|H_i)$ будет обозначать функцию правдоподобия при гипотезе H_i $(i = 0, 1)$ и будет использоваться только один знак интеграла для обозначения n-кратного интегрирования по выборочному пространству. Задача состоит в выборе критической области w, максимизирующей выражение (22.2), которое в интегральной записи имеет вид

$$1 - \beta = \int_w L(x|H_1)\, dx \tag{22.3}$$

при условии (22.1), которое мы запишем теперь в виде

$$\int_w L(x|H_0)\, dx = \alpha. \tag{22.4}$$

Эта критическая область, очевидно, должна включать все точки x, для которых $L(x|H_0) = 0$, $L(x|H_1) > 0$, так как эти точки не дают никакого вклада в интеграл (22.4). Для остальных точек области w можно переписать (22.3) в виде

$$1 - \beta = \int_w \frac{L(x|H_1)}{L(x|H_0)} L(x|H_0)\, dx. \tag{22.5}$$

Таким образом, надо выбрать w так, чтобы максимизировать математическое ожидание отношения $\frac{L(x|H_1)}{L(x|H_0)}$, вычисленное в предположении, что справедлива гипотеза H_0. Ясно, что это будет выполнено тогда и только тогда, когда w, удовлетворяя условию (22.4), содержит точки, для которых отношение $\frac{L(x|H_1)}{L(x|H_0)}$ принимает наибольшие значения. Таким образом, НКО состоит из тех точек пространства W, для которых

$$\frac{L(x|H_0)}{L(x|H_1)} \leqslant k_\alpha. \tag{22.6}$$

Каждой константе k_α в (22.6) соответствует размер α, определяемый (22.4). Если иксы распределены непрерывно, то мы можем также найти для каждого α соответствующее k_α.

22.11 Если распределение иксов не непрерывно, то можно использовать прием рандомизации (см. **20.22**). В этом случае

$$\frac{L(x \mid H_0)}{L(x \mid H_1)} = k_\alpha \tag{22.7}$$

с некоторой ненулевой вероятностью p, и в общем случае из-за дискретности можно выбрать в (22.6) лишь такое k_α, которому соответствует размер критерия $\alpha - q$ $(0 < q < p)$. Чтобы получить критерий, размер которого точно равен α, надо поступить следующим образом. Всякий раз, когда выполняется (22.7), используется случайное устройство (например, таблица случайных чисел) с тем, чтобы с вероятностью q/p гипотеза H_0 отвергалась и с вероятностью $1 - (q/p)$ принималась. Полная вероятность отвергнуть H_0 будет тогда равна $(\alpha - q) + pq/p = \alpha$, как это и требовалось. В этом случае, очевидно, НКО не единственна, поскольку она подвержена случайным выборочным флуктуациям *).

Пример 22.2

Вернемся снова к нормальному распределению примера 22.1

$$dF(x) = (2\pi)^{-1/2} \exp\left[-\tfrac{1}{2}(x-\mu)^2\right] dx, \quad -\infty \leqslant x \leqslant \infty, \tag{22.8}$$

и будем проверять гипотезу H_0: $\mu = \mu_0$ против альтернативы

$$H_1\colon\ \mu = \mu_1 \qquad (\neq \mu_0).$$

Имеем

$$L(x \mid H_i) = (2\pi)^{-n/2} \exp\left[-\frac{1}{2}\sum_{j=1}^{n}(x_j - \mu_i)^2\right] =$$

$$= (2\pi)^{-n/2} \exp\left[-\frac{n}{2}\{s^2 + (\bar{x} - \mu_i)^2\}\right], \quad i = 0, 1, \tag{22.9}$$

*) При применении рандомизации не имеет смысла говорить о критической области. Критерий в этом случае задается *критической функцией* $\varphi(x)$ $(0 \leqslant \varphi(x) \leqslant 1)$, дающей вероятность отвергнуть H_0 при наблюдении x. Например, критерию в тексте отвечает критическая функция $\varphi(x) = 1$, q/p или 0, когда $L(x|H_0)/L(x|H_1) < k_\alpha$, $= k_\alpha$ или $> k_\alpha$ соответственно. Отсутствию рандомизации отвечает критическая функция, принимающая два значения: 1 на ω и 0 в остальных точках. (*Прим. ред.*)

где $\bar{x}$, s^2 — выборочные среднее и дисперсия. Следовательно, соотношение (22.6), определяющее НКО, записывается в виде

$$\frac{L(x \mid H_0)}{L(x \mid H_1)} = \exp\left[\frac{n}{2}\{(\bar{x}-\mu_1)^2 - (\bar{x}-\mu_0)^2\}\right] =$$

$$= \exp\left[\frac{n}{2}\{(\mu_0-\mu_1)2\bar{x} + (\mu_1^2 - \mu_0^2)\}\right] \leqslant k_\alpha \qquad (22.10)$$

или

$$(\mu_0 - \mu_1)\bar{x} \leqslant \frac{1}{2}(\mu_0^2 - \mu_1^2) + \frac{1}{n}\log k_\alpha. \qquad (22.11)$$

Таким образом, при заданных μ_0, μ_1 и α НКО определяется только значением выборочного среднего $\bar{x}$. Этого следовало ожидать, принимая во внимание тот факт (см. примеры 17.6, 17.15), что $\bar{x}$ является достаточной статистикой для μ, соответствующей минимальной границе дисперсии. Далее, из (22.11) следует, что если $\mu_0 > \mu_1$, то НКО имеет вид

$$\bar{x} \leqslant \frac{1}{2}(\mu_0 + \mu_1) + \log k_\alpha/\{n(\mu_0 - \mu_1)\}, \qquad (22.12)$$

тогда как при $\mu_0 < \mu_1$ мы получаем

$$\bar{x} \geqslant \frac{1}{2}(\mu_0 + \mu_1) - \log k_\alpha/\{n(\mu_1 - \mu_0)\}, \qquad (22.13)$$

что вполне согласуется с интуицией: проверяя гипотетическое значение μ_0 против меньшего значения μ_1, мы отвергаем μ_0, если выборочное среднее оказывается меньше некоторого значения, зависящего от размера критерия α; проверяя μ_0 против большего значения μ_1, мы отвергаем μ_0, если выборочное среднее превосходит некоторое определенное значение.

22.12 Следует отметить, что в примере 22.2 для определения НКО требуется знание только одной статистики, а не всей совокупности наблюдений. Это характерно для многих задач. Такое упрощение позволит нам вести изложение целиком в терминах выборочного распределения такой статистики (называемой «статистикой критерия») и избежать осложнений, связанных с рассмотрением n-мерных распределений.

Пример 22.3

Известно (см. пример 11.12), что в примере 22.2 при любом значении μ выборочное среднее $\bar{x}$ само распределено нормально со средним μ и дисперсией $1/n$. Таким образом, чтобы получить критерий размера α для проверки μ_0 против $\mu_1 > \mu_0$, мы должны найти такое $\bar{x}_\alpha$, чтобы выполнялось равенство

$$\int_{\bar{x}_\alpha}^{\infty} \left(\frac{n}{2\pi}\right)^{1/2} \exp\left\{-\frac{n}{2}(\bar{x}-\mu_0)^2\right\} d\bar{x} = \alpha.$$

Обозначая

$$G(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} \exp\left(-\frac{1}{2}y^2\right) dy, \tag{22.14}$$

получаем для $\mu_1 > \mu_0$

$$\bar{x}_\alpha = \mu_0 + d_\alpha/n^{1/2}, \tag{22.15}$$

где

$$G(-d_\alpha) = \alpha. \tag{22.16}$$

Например, для $\mu_0 = 2$, $n = 25$ и $\alpha = 0{,}05$ из таблицы нормального интеграла находим

$$d_{0,05} = 1{,}6449,$$

так что согласно (22.15)

$$\bar{x}_\alpha = 2 + 1{,}6449/5 = 2{,}3290.$$

В этом простом примере мощность критерия может быть выписана в явном виде. Она равна

$$\int_{\bar{x}_\alpha}^{\infty} \left(\frac{n}{2\pi}\right)^{1/2} \exp\left\{-\frac{n}{2}(\bar{x} - \mu_1)^2\right\} d\bar{x} = 1 - \beta. \tag{22.17}$$

Используя (22.15), можно представить этот интеграл в виде

$$1 - G\{n^{1/2}(\mu_0 - \mu_1) + d_\alpha\} = G\{n^{1/2}(\mu_1 - \mu_0) - d_\alpha\}, \tag{22.18}$$

поскольку $G(x) = 1 - G(-x)$ в силу симметрии. Из (22.18) видно, что мощность является монотонно возрастающей функцией как по n (объем выборки), так и по $\mu_1 - \mu_0$ (разность гипотетических значений, между которыми делается выбор).

Пример 22.4

Для противопоставления рассмотрим распределение Коши

$$dF(x) = \frac{dx}{\pi\{1 + (x - \theta)^2\}}, \qquad -\infty \leqslant x \leqslant \infty,$$

и предположим, что мы хотим проверить гипотезу H_0: $\theta = 0$ против альтернативы H_1: $\theta = 1$. Для простоты ограничимся случаем $n = 1$. Согласно (22.6) НКО определяется соотношением

$$\frac{L(x \mid H_0)}{L(x \mid H_1)} = \frac{1 + (x - 1)^2}{1 + x^2} \leqslant k_\alpha,$$

которое эквивалентно соотношению

$$x^2(k_\alpha - 1) + 2x + (k_\alpha - 2) \geqslant 0. \tag{22.19}$$

Вид НКО зависит, таким образом, от выбранного значения α. Если, например, $k_\alpha = 1$, то (22.19) сводится к неравенству $x \geqslant 1/2$, т. е. мы должны отвергнуть $\theta = 0$ в пользу $\theta = 1$ каждый раз, когда наблюденная величина ближе к 1, чем к 0. Если, с другой стороны, взять $k_\alpha = 0{,}5$, то (22.19) принимает вид

$$x^2 - 4x + 3 \leqslant 0 \quad \text{или} \quad (x-2)^2 \leqslant 1.$$

Последнее неравенство выполнено, когда $1 \leqslant x \leqslant 3$. Это и есть критическая область.

Поскольку распределение Коши есть распределение Стьюдента с одной степенью свободы и, следовательно, $F(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} x$, то можно вычислить размер каждого из двух приведенных выше критериев.

Для $k_\alpha = 1$ размер равен $\mathbf{P}\left(t \geqslant \frac{1}{2}\right) = 0{,}352$, тогда как для $k_\alpha = 0{,}5$ получаем

$$\mathbf{P}\,(1 \leqslant t \leqslant 3) = 0{,}148.$$

Использованная таблица может быть применена также для нахождения мощности рассмотренных критериев. Это предлагается проделать читателю в упражнении 22.4 в конце этой главы.

22.13 Приведенные нами примеры использования леммы Неймана — Пирсона касались проверки гипотезы о значении параметра некоторого распределения заданного вида. Однако, как можно видеть из доказательства в **22.10—11**, (22.6) дает НКО для любой задачи проверки простой гипотезы против простой альтернативы. Например, можно проверить гипотезу о виде распределения, параметр расположения которого известен, как это делается в следующем примере.

Пример 22.5

Пусть известно, что среднее значение распределения равно нулю и нас интересует вид этого распределения. Предположим, что мы хотим выбрать одну из альтернатив

$$\left.\begin{aligned} H_0\colon\ & dF = (2\pi)^{-1/2}\exp\left(-\tfrac{1}{2}x^2\right)dx, \\ H_1\colon\ & dF = \tfrac{1}{2}\exp\left(-|x|\right)dx, \end{aligned}\right\} \quad -\infty \leqslant x \leqslant \infty.$$

Для простоты мы опять будем считать объем выборки n равным 1.

Согласно (22.6) НКО определяется соотношением

$$\frac{L(x \mid H_0)}{L(x \mid H_1)} = \left(\frac{2}{\pi}\right)^{1/2} \exp\left(|x| - \frac{1}{2}x^2\right) \leqslant k_\alpha.$$

Таким образом, мы отвергаем H_0, если

$$|x| - \frac{1}{2}x^2 \leqslant \log\left\{k_\alpha \left(\frac{\pi}{2}\right)^{1/2}\right\} = c_\alpha.$$

НКО состоит, следовательно, из больших по абсолютной величине значений наблюдения, дополненных, если $k_\alpha > (2/\pi)^{1/2}$ (т. е. $c_\alpha > 0$), значениями из окрестности $x = 0$. Читатель может легко убедиться в этом, построив график.

НКО и достаточные статистики

22.14 Если обе сравниваемые гипотезы относятся к значению параметра θ, для которого имеется достаточная статистика t, то из факторизации функции правдоподобия, выражаемой формулой (17.18), следует, что (22.6) имеет в этом случае вид

$$\frac{L(x \mid \theta_0)}{L(x \mid \theta_1)} = \frac{g(t \mid \theta_0)}{g(t \mid \theta_1)} \leqslant k_\alpha, \tag{22.20}$$

т. е., как этого и следовало ожидать, НКО зависит от значения достаточной статистики t. Мы уже встречались с таким случаем в примере 22.2. (Такой же результат, очевидно, имеет место, если θ представляет собой набор параметров, для которого t служит достаточной системой статистик.) Упражнение 22.13 показывает, что отношение правдоподобия в левой части (22.10) само является достаточной статистикой, так что НКО есть функция от этой достаточной статистики.

Однако не всегда получается так, что НКО имеет вид $t \geqslant a_\alpha$ или $t \leqslant b_\alpha$, как это было в примере 22.2 (контрпримером служит пример 22.4, в котором единственное наблюдение x является достаточной статистикой для θ). Исследование выражения (22.20) показывает, что НКО будет иметь такой простой вид, если $g(t|\theta_0)/g(t|\theta_1)$ является монотонно неубывающей функцией от t для $\theta_0 > \theta_1$. Это будет заведомо выполнено, если

$$\frac{\partial^2}{\partial\theta\,\partial t} \log g(t \mid \theta) \geqslant 0, \tag{22.21}$$

что справедливо почти для всех распределений, встречающихся в статистике.

Пример 22.6

В случае распределения

$$dF(x) = \begin{cases} \exp\{-(x-\theta)\}\,dx, & \theta \leqslant x \leqslant \infty, \\ 0 \text{ в противном случае} \end{cases}$$

наименьшее выборочное значение $x_{(1)}$ достаточно для θ (см. пример 17.19). Проверяя гипотезу θ_0 против $\theta_1 > \theta_0$ для выборки из n наблюдений, мы получим

$$\frac{L(x \mid \theta_0)}{L(x \mid \theta_1)} = \begin{cases} \infty, & \text{если} \quad x_{(1)} < \theta_1, \\ \exp\{n(\theta_0 - \theta_1)\} & \text{в противном случае,} \end{cases}$$

откуда следует, что для точек критической области должно быть выполнено неравенство

$$\exp\{n(\theta_0 - \theta_1)\} \leqslant k_\alpha. \tag{22.22}$$

Поскольку, однако, левая часть (22.22) вообще не зависит от наблюдений и является константой, то (22.22) будет выполнено для любой критической области размера α с $x_{(1)} \geqslant \theta_1$. Таким образом, все такие критические области имеют одинаковую мощность и будут, следовательно, НКО.

Если допустить, что θ_1 может быть как больше, так и меньше θ_0, то мы получим

$$\frac{L(x \mid \theta_0)}{L(x \mid \theta_1)} = \begin{cases} \infty, & \text{если} \quad \theta_0 \leqslant x_{(1)} < \theta_1, \\ \exp\{n(\theta_0 - \theta_1)\} > 1, & \text{если} \quad x_{(1)} \geqslant \theta_0 > \theta_1, \\ \exp\{n(\theta_0 - \theta_1)\} < 1, & \text{если} \quad x_{(1)} \geqslant \theta_1 > \theta_0, \\ 0, & \text{если} \quad \theta_1 \leqslant x_{(1)} < \theta_0. \end{cases}$$

Таким образом, НКО определяется неравенствами

$$(x_{(1)} - \theta_0) < 0, \qquad (x_{(1)} - \theta_0) > c_\alpha.$$

Первое из этих событий имеет при гипотезе H_0 вероятность нуль. Значение c_α выбирается так, чтобы вероятность осуществления второго события при гипотезе H_0 была равна α.

Эффективность оценки и мощность

22.15 Если для проверки гипотезы используется статистика, эффективная в задаче оценивания (см. **17.28—29**), то из этого не следует, что мощность полученного критерия будет больше мощности критерия, основанного на менее эффективной оценке. Этот результат принадлежит Сандруму (1954).

Пусть t_1 и t_2 — две асимптотически нормально распределенные оценки параметра θ, и пусть, по крайней мере асимптотически,

$$\left.\begin{aligned} \mathsf{M}(t_1) &= \mathsf{M}(t_2) = \theta, \\ \mathsf{D}(t_i \mid \theta = \theta_0) &= \sigma_{i0}^2, \\ \mathsf{D}(t_i \mid \theta = \theta_1) &= \sigma_{i1}^2, \end{aligned}\right\} \quad i = 1, 2.$$

Пусть проверяется гипотеза H_0: $\theta = \theta_0$ против H_1: $\theta = \theta_1 > \theta_0$. Точно так же, как в (22.15) в примере 22.3, мы для каждой из статистик получим критическую область

$$t_i \geqslant \theta_0 + d_\alpha \sigma_{i0}, \qquad i = 1, 2, \tag{22.23}$$

где d_α — нормальное отклонение, определяемое соотношениями (22.14) и (22.16). Обобщая равенство (22.18), соответствующее случаю $\sigma_{i0} = \sigma_{i1}$, получаем мощности критериев

$$1 - \beta(t_i) = G\left\{\frac{(\theta_1 - \theta_0) - d_\alpha \sigma_{i0}}{\sigma_{i1}}\right\}. \tag{22.24}$$

Поскольку $G(x)$ — монотонно возрастающая функция своего аргумента, то в соответствии с (22.24) t_1 приводит к более мощному критерию, чем t_2, тогда и только тогда, когда

$$\frac{(\theta_1 - \theta_0) - d_\alpha \sigma_{10}}{\sigma_{11}} > \frac{(\theta_1 - \theta_0) - d_\alpha \sigma_{20}}{\sigma_{21}},$$

т. е. когда

$$\theta_1 - \theta_0 > d_\alpha \left(\frac{\sigma_{10}\sigma_{21} - \sigma_{20}\sigma_{11}}{\sigma_{21} - \sigma_{11}}\right). \tag{22.25}$$

Если положить $E_j = \sigma_{2j}/\sigma_{1j}$ $(j = 0, 1)$, то (22.25) примет вид

$$\theta_1 - \theta_0 > d_\alpha \left(\frac{E_1 - E_0}{E_1 - 1}\right) \sigma_{10}. \tag{22.26}$$

Величины E_0 и E_1 представляют собой просто степени (обычно квадратные корни) эффективности t_1 относительно t_2 при выполнении соответственно гипотез H_0 и H_0. Если теперь

$$E_0 = E_1 > 1, \tag{22.27}$$

то правая часть (22.25) обращается в нуль и (22.26) всегда выполняется. Таким образом, если относительная эффективность оценок t_1 и t_2 *одна и та же* при обеих гипотезах, то более эффективная статистика t_1 всегда приводит к более мощному критерию, независимо от значений α или $\theta_1 - \theta_0$. Однако если

$$E_1 > E_0 \geqslant 1, \tag{22.28}$$

то всегда можно найти такое достаточно малое α, что для критерия этого размера (22.26) не будет выполнено. Следовательно, критерий, основанный на менее эффективной оценке t_2, будет иметь бóльшую мощность, если выполнено соотношение (22.28), т. е. если относительная эффективность t_2 при гипотезе H_0 больше, чем при H_1. Если же $E_0 > E_1 > 1$, то можно найти такое достаточно большое α, чтобы (22.26) было нарушено. Если функция E_1 непрерывна по θ, то $E_1 \to E_0$ при $\theta_1 \to \theta_0$, так что (22.26) выполняется в некоторой окрестности θ_0.

Приведенный результат хотя и является довольно частным, все же с достаточной ясностью показывает, что между эффек-

тивностью при оценивании и мощностью нет тесной связи. В главе 25 мы снова вернемся к этому вопросу в связи с рассмотрением способов измерения эффективности критериев.

Пример 22.7

В примерах 18.3 и 18.6 было показано, что МП-оценка $\hat{\rho}$ параметра ρ нормированного двумерного нормального распределения является корнем некоторого кубического уравнения и что асимптотическая дисперсия этой оценки равна $(1-\rho^2)^2/\{n(1+\rho^2)\}$, тогда как выборочный коэффициент корреляции r имеет асимптотическую дисперсию $(1-\rho^2)^2/n$. Обе оценки состоятельны и асимптотически нормальны, а МП-оценка, кроме того, эффективна. В обозначениях пункта **22.15** имеем

$$E=(1+\rho^2)^{1/2}.$$

Если проверяется гипотеза H_0: $\rho=0$ против H_1: $\rho=0{,}1$, то $E_0=1$ и (22.26) принимает вид

$$0{,}1>d_\alpha\sigma_{10}=d_\alpha\,(1/n)^{1/2}. \qquad (22.29)$$

Если взять, скажем, $n=400$, чтобы было справедливо нормальное приближение, то (22.29) будет неверно при $d_\alpha>2$. Это соответствует $\alpha<0{,}023$, так что для критериев размера, меньшего 0,023, неэффективная оценка r дает в нашем случае критерий с большей асимптотической мощностью, чем эффективная оценка $\hat{\rho}$. Поскольку критерии размера 0,01 и 0,05 используются довольно часто, то этот пример представляет не только теоретический интерес: на практике нельзя предполагать, что «хорошие» оценки являются «хорошими» статистиками критериев.

Проверка простой гипотезы H_0 против класса альтернатив

22.16 До сих пор мы рассматривали наиболее элементарную задачу, в которой требовалось сделать выбор между двумя полностью определенными конкурирующими гипотезами. В такой задаче имеется очевидная симметрия: какая из гипотез считается «проверяемой» и какая «альтернативной», зависит только от соглашения или от удобства. При переходе к более общей формулировке задачи проверки гипотез симметрия исчезает.

Мы будем рассматривать теперь случай, когда H_0 — простая гипотеза, а H_1 — сложная гипотеза, представляющая собой класс простых альтернатив. Наиболее часто встречается случай, когда имеется класс Ω простых параметрических гипотез, одна из которых — гипотеза H_0, а H_1 включает все остальные. Например, гипотеза H_0 может состоять в том, что среднее значение

некоторого распределения равно μ_0, а гипотеза H_1 — в том, что среднее не равно μ_0.

Для каждой простой гипотезы H_t из H_1 можно применить полученные выше результаты и найти для заданного α соответствующую ей НКО w_t. Однако, вообще говоря, эти области будут разными для разных H_t. Нас, конечно, не устраивает такая ситуация, поэтому возникает вопрос о существовании одной НКО, которая была бы наилучшей для всех H_t из H_1. Такая область называется равномерно наиболее мощной (РНМ), а критерий, основанный на этой области, — РНМ критерием.

22.17 К сожалению, как мы увидим в дальнейшем, РНМ критерий обычно не существует, если только класс альтернатив Ω не удовлетворяет определенным ограничениям. Рассмотрим, например, случай, с которым мы имели дело в примере 22.2. Мы нашли, что при $\mu_1 < \mu_0$ НКО для простой альтернативы определяется неравенством

$$\bar{x} \leqslant a_\alpha. \tag{22.30}$$

Мы видим, что при $\mu_1 < \mu_0$ области, определяемые соотношением (22.30), не зависят от μ_1 и могут быть найдены непосредственно из выборочного распределения $\bar{x}$, если задан размер критерия α. Следовательно, критерий, основанный на (22.30), будет РНМ критерием для класса гипотез $\mu_1 < \mu_0$.

Из примера 22.2 также следует, что для $\mu_1 > \mu_0$ НКО определяется неравенством $\bar{x} \geqslant b_\alpha$. Так что если наш класс Ω состоит из значений μ_1, больших μ_0, то этот критерий будет РНМ. Однако если μ_1 может быть как больше, так и меньше μ_0, то РНМ критерия не существует, так как всегда одна из двух только что найденных РНМ областей будет лучше любой компромиссной области.

22.18 Мы докажем теперь, что для простой гипотезы H_0: $\theta = \theta_0$, относящейся к параметру θ, порождающему класс гипотез, в общем случае нельзя найти РНМ критерия, если класс альтернатив является интервалом, содержащим как положительные, так и отрицательные значения $\theta - \theta_0$. При этом предполагаются выполненными некоторые условия регулярности, в частности, производная функции правдоподобия по θ предполагается непрерывной по θ.

Разложим функцию правдоподобия в ряд Тейлора в окрестности θ_0:

$$L(x \mid \theta_1) = L(x \mid \theta_0) + (\theta_1 - \theta_0) L'(x \mid \theta^*), \tag{22.31}$$

где θ^* — некоторое значение из интервала (θ_1, θ_0). Тогда для НКО, если она существует, мы должны иметь согласно (22.6) и (22.31)

$$\frac{L(x \mid \theta_1)}{L(x \mid \theta_0)} = 1 + \frac{(\theta_1 - \theta_0) L'(x \mid \theta^*)}{L(x \mid \theta_0)} \geqslant k_\alpha(\theta_1). \tag{22.32}$$

Таким образом, НКО определяется соотношениями

$$\frac{L'(x\mid\theta^*)}{L(x\mid\theta_0)} \geqslant a_\alpha, \quad \theta_1 > \theta_0, \tag{22.33}$$

$$\frac{L'(x\mid\theta^*)}{L(x\mid\theta_0)} \leqslant b_\alpha, \quad \theta_1 < \theta_0. \tag{22.34}$$

Посмотрим теперь, что происходит, когда θ_1 приближается к θ_0. Очевидно, θ^* тогда тоже будет приближаться к θ_0, и для θ_1, очень близких к θ_0, неравенства (22.33), (22.34) ввиду непрерывности L' по θ примут вид

$$\frac{L'(x\mid\theta_0)}{L(x\mid\theta_0)} = \left[\frac{\partial \log L}{\partial\theta}\right]_{\theta=\theta_0} \geqslant a_\alpha, \quad \theta > \theta_0, \tag{22.35}$$

$$\frac{L'(x\mid\theta_0)}{L(x\mid\theta_0)} = \left[\frac{\partial \log L}{\partial\theta}\right]_{\theta=\theta_0} \leqslant b_\alpha, \quad \theta < \theta_0. \tag{22.36}$$

Тем самым мы установили, между прочим, что в близкой окрестности θ_0 односторонние критерии, основанные на $\left[\frac{\partial \log L}{\partial\theta}\right]_{\theta=\theta_0}$, будут РНМ. Это является аналогом результата из теории доверительных интервалов, полученного в **20.17**.

Наш основной результат получается теперь немедленно. Если мы рассмотрим интервал альтернатив, включающий положительные и отрицательные значения $(\theta_1 - \theta_0)$, то (22.35) и (22.36) не смогут выполняться одновременно (и, следовательно, не может существовать НКО), за исключением случая, когда

$$\left[\frac{\partial \log L}{\partial\theta}\right]_{\theta=\theta_0} = \text{const}. \tag{22.37}$$

Условие (22.37) является существенным условием существования двусторонней НКО. Оно не может быть выполнено, если выполнено условие (17.18) (т. е. область определения распределения не зависит от θ), кроме случая, когда константа равна нулю, так как условие $\mathbf{M}\left[\frac{\partial \log L}{\partial\theta}\right] = 0$ вместе с (22.37) дает $\left[\frac{\partial \log L}{\partial\theta}\right]_{\theta=\theta_0} = 0$.

В примере 22.6 мы уже встречались со случаем существования двусторонней НКО. Читателю следует проверить, что в этом примере $\left[\frac{\partial \log L}{\partial\theta}\right]_{\theta=\theta_0} = n$, так что (22.37) выполнено.

РНМ критерии в случае нескольких параметров

22.19 Если рассматриваемое распределение имеет более одного параметра и проверяется простая гипотеза, то в этом случае также может существовать общая НКО для класса альтернативных гипотез, получающихся при различных значениях

параметров. В следующих двух примерах рассматривается случай двухпараметрического нормального распределения, для которого НКО не существует, и случай двухпараметрического экспоненциального распределения, для которого НКО существует.

Пример 22.8

Рассмотрим нормальное распределение со средним μ и дисперсией σ^2. Пусть проверяется гипотеза

$$H_0\colon\ \mu = \mu_0, \qquad \sigma = \sigma_0,$$

а от альтернативной гипотезы H_1 требуется лишь отличие от H_0. Для любой простой гипотезы

$$H_1\colon\ \mu = \mu_1, \qquad \sigma = \sigma_1,$$

НКО согласно (22.6) определяется неравенством

$$\frac{L(x\,|\,H_0)}{L(x\,|\,H_1)} = \left(\frac{\sigma_1}{\sigma_0}\right)^n \exp\left[-\frac{1}{2}\left\{\sum\left(\frac{x-\mu_0}{\sigma_0}\right)^2 - \sum\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right\}\right] \leqslant k_\alpha,$$

которое может быть переписано в виде

$$\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2}\right) + \frac{(\bar{x}-\mu_1)^2}{\sigma_1^2} - \frac{(\bar{x}-\mu_0)^2}{\sigma_0^2} \leqslant \frac{2}{n}\log\left\{\left(\frac{\sigma_0}{\sigma_1}\right)^n k_\alpha\right\},$$

где $\bar{x}$ и s^2 — выборочные среднее и дисперсия. Если $\sigma_0 \neq \sigma_1$, то можно привести это соотношение к еще более простому виду:

$$\left(\frac{\sigma_1^2 - \sigma_0^2}{\sigma_0^2\sigma_1^2}\right)\sum(x-\rho)^2 \geqslant c_\alpha, \tag{22.38}$$

где c_α не зависит от наблюдений и

$$\rho = \frac{\mu_0\sigma_1^2 - \mu_1\sigma_0^2}{\sigma_1^2 - \sigma_0^2}.$$

Со случаем $\sigma_0 = \sigma_1$ мы уже имели дело в примере 22.2, где предполагалось, что $\sigma_0 = \sigma_1 = 1$.

Когда в (22.38) имеет место точное равенство, оно превращается в уравнение гиперсферы с центром в точке с координатами $x_1 = \ldots = x_n = \rho$. Таким образом, НКО всегда ограничена гиперсферой. При $\sigma_1 > \sigma_0$ (22.38) дает $\sum(x-\rho)^2 \geqslant a_\alpha$, так что НКО лежит вне сферы, а при $\sigma_1 < \sigma_0$ из (22.38) находим $\sum(x-\rho)^2 \leqslant b_\alpha$, т. е. НКО лежит внутри сферы.

Поскольку ρ есть функция от μ_1 и σ_1, то ясно, что в общем случае мы не получим одной и той же НКО для различных гипотез из H_1, даже если предположим, что $\sigma_1 < \sigma_0$ и $\mu_1 < \mu_0$

или наложим другие подобные ограничения. Для иллюстрации воспользуемся графиком на плоскости $(\bar{x}, s)$. Так как

$$\sum (x-\rho)^2 = \sum (x-\bar{x})^2 + n(\bar{x}-\rho)^2 = n\{s^2 + (\bar{x}-\rho)^2\}, \quad (22.39)$$

то, когда (22.39) постоянно, мы получим на графике круг с центром $(\rho, 0)$ и фиксированным радиусом, зависящим от α.

На рис. 22.2 (Нейман и Пирсон, 1933b) изображены некоторые кривые, соответствующие различным частным случаям. В каждом случае изображена одна кривая, соответствующая некоторому фиксированному значению k, константы в (22.37).

Случаи (1) и (2): $\sigma_1 = \sigma_0$ и $\rho = \pm\infty$. НКО лежит справа от прямой (1), если $\mu_1 > \mu_0$, и слева от прямой (2), если $\mu_1 < \mu_0$. Этот случай рассмотрен в примере 22.2.

Случай (3): $\sigma_1 < \sigma_0$; пусть, например, $\sigma_1 = \frac{1}{2}\sigma_0$. Тогда $\rho = \mu_0 + \frac{4}{3}(\mu_1 - \mu_0)$ и НКО лежит внутри полукруга (3).

Случай (4): $\sigma_1 < \sigma_0$ и $\mu_1 = \mu_0$. НКО лежит внутри полукруга (4).

Случай (5): $\sigma_1 > \sigma_0$ и $\mu_1 = \mu_0$. НКО лежит вне полукруга (5).

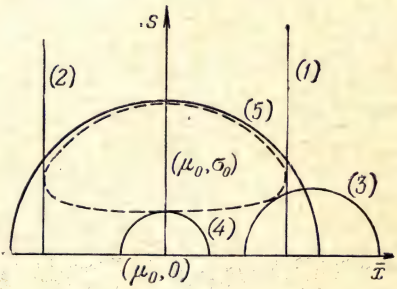


Рис. 22.2. Линии постоянного отношения правдоподобия k (см. текст).

Очевидно, для этих случаев не имеется общей НКО. Области принятия гипотезы, однако, могут иметь общую часть, расположенную около точки (μ_0, σ_0), и этого естественно было ожидать. Найдем огибающую НКО, которая, конечно, совпадает с огибающей областей принятия. Продифференцируем отношение правдоподобия по μ_1 и σ_1 и приравняем полученные производные нулю. Это даст решения уравнений МП (см. пример 18.11) $\hat{\mu} = \bar{x}$, $\hat{\sigma}_1 = s$. Подставляя их в отношение правдоподобия, мы найдем уравнение огибающей

$$1 - \frac{s^2}{\sigma_0^2} - \left(\frac{\bar{x}-\mu_0}{\sigma_0}\right)^2 = \frac{2}{n}\log\left\{\left(\frac{\sigma_0}{s}\right)^n k_\alpha\right\},$$

или

$$\left(\frac{\bar{x}-\mu_0}{\sigma_0}\right)^2 - \log\left(\frac{s^2}{\sigma_0^2}\right) + \frac{s^2}{\sigma_0^2} = 1 - \frac{2}{n}\log k_\alpha. \quad (22.40)$$

Пунктирная кривая на рис. 22.2 изображает огибающую. Она касается границ всех НКО, которые отвечают одному и тому же k (и, следовательно, имеют разные размеры α). Часть плоскости, заключенная внутри кривой, может рассматриваться как

«хорошая» область принятия гипотезы, а внешняя часть — как хорошая критическая область. В данном случае, не существует НКО для всех альтернатив, однако области, определяемые огибающими областей отношения правдоподобия, дают компромиссное решение, выделяя и объединяя части критических областей, являющиеся наилучшими для отдельных альтернатив.

Пример 22.9

Пусть требуется проверить простую гипотезу H_0: $\theta = \theta_0$, $\sigma = \sigma_0$, против альтернативы H_1: $\theta = \theta_1 \leqslant \theta_0$, $\sigma = \sigma_1 < \sigma_0$, в случае распределения

$$dF = \exp\left\{-\left(\frac{x-\theta}{\sigma}\right)\right\}\frac{dx}{\sigma}, \qquad \theta \leqslant x \leqslant \infty, \qquad \sigma > 0. \tag{22.41}$$

Согласно (22.6) НКО определяется неравенством

$$\frac{L_0}{L_1} = \left(\frac{\sigma_1}{\sigma_0}\right)^n \exp\left\{-\frac{n(\bar{x}-\theta_0)}{\sigma_0} + \frac{n(\bar{x}-\theta_1)}{\sigma_1}\right\} \leqslant k_\alpha,$$

так что для любых значений θ_1, σ_1 из H_1 НКО имеет вид

$$x_{(1)} \leqslant \theta_0, \quad \bar{x} \leqslant \frac{\frac{1}{n}\log\left\{k_\alpha\left(\frac{\sigma_0}{\sigma_1}\right)^n\right\} + \left(\frac{\theta_1}{\sigma_1} - \frac{\theta_0}{\sigma_0}\right)}{\left(\frac{1}{\sigma_1} - \frac{1}{\sigma_0}\right)}. \tag{22.42}$$

Когда выполнена гипотеза H_0, первое из этих событий имеет вероятность нуль. Следовательно, для всего класса альтернатив H_1 существует одна и та же НКО, на которой может быть основан РНМ критерий.

Со случаем $\sigma_1 = \sigma_0$ мы фактически имели дело в примере 22.6.

РНМ критерии и достаточные статистики

22.20 В **22.14** было показано, что при проверке простой параметрической гипотезы против простой альтернативы НКО определяется (совместно) достаточной статистикой для параметра (параметров), если такая существует. Если же проверяется простая гипотеза H_0 против сложной гипотезы H_1, представляющей собой класс простых параметрических альтернатив, то, как непосредственно вытекает из рассуждений пункта **22.14**, если существует общая НКО, дающая РНМ критерий, и t — достаточная статистика для параметра (параметров), то НКО будет определяться статистикой t. Однако поскольку РНМ критерий не всегда существует, то возникают новые вопросы. Следует ли из существования РНМ критерия существование соответствую-

щей достаточной статистики? И наоборот, гарантирует ли существование достаточной статистики существование соответствующего РНМ критерия?

22.21 На первый из вопросов может быть дан утвердительный ответ при одном дополнительном ограничении. Как показали Нейман и Пирсон (1936а), если (1) существует общая НКО и, следовательно, существует РНМ критерий для H_0 против H_1 любого размера α в интервале $0 < \alpha \leqslant \alpha_0$ (где α_0 не обязательно равно 1) и если (2) каждая точка выборочного пространства W (за исключением, возможно, точек множества меры нуль) принадлежит границе НКО по крайней мере для одного значения α и при этом соответствует значению $L(x|H_0) > 0$, то существует одномерная достаточная статистика для параметра (параметров), изменение которого (которых) порождает класс допустимых альтернативных гипотез H_1.

Чтобы установить этот результат, мы заметим сначала, что если НКО, общая для всех альтернатив класса H_1, существует для двух размеров критериев α_1 и $\alpha_2 < \alpha_1$, то общая НКО размера α_2 может быть всегда получена как подобласть размера α_1. Это следует из того факта, что любая общая НКО удовлетворяет (22.6). Мы можем, следовательно, не теряя общности, считать, что, когда α убывает, соответствующие НКО получаются просто исключением некоторых точек *).

Предположим теперь, что выполнены условия (1) и (2). Если некоторая точка (скажем, x) области W принадлежит границе НКО только при одном значении α, то мы определим статистику t равенством

$$t(x) = \alpha. \tag{22.43}$$

Если точка x принадлежит границе НКО при более чем одном значении α, то, обозначая α_1 и α_2 наименьшее и наибольшее из этих значений, положим

$$t(x) = \frac{1}{2}(\alpha_1 + \alpha_2). \tag{22.44}$$

Из замечания в предыдущем абзаце следует, что x будет принадлежать границе НКО для всех α из интервала (α_1, α_2). Таким образом, для всех точек W (за исключением, возможно, множества меры нуль) соотношениями (22.43) и (22.44) определена статистика t. Далее, если t принимает одно и то же значение в двух точках, то они должны принадлежать одной и той же границе. Следовательно, из (22.6) мы получаем

$$\frac{L(x|\theta_0)}{L(x|\theta)} = k(t, \theta),$$

*) Для критических областей в общем случае это неверно. См., например, Чернов (1951).

где k зависит от наблюдений только через статистику t. Таким образом, мы должны иметь

$$L(x|\theta) = g(t|\theta)\, h(x), \tag{22.45}$$

так что одномерная статистика t достаточна для рассматриваемого набора параметров θ.

22.22 Мы уже рассматривали в примере 22.2 случай, когда существует одновременно одномерная достаточная статистика и РНМ критерий. В упражнениях 22.1—22.3 содержатся дальнейшие примеры. Однако если условие (2) пункта **22.21** не выполнено, то из существования РНМ критерия может не следовать существование одномерной достаточной статистики. Следующий пример иллюстрирует это утверждение.

Пример 22.10

В примере 22.9 было показано, что распределение (22.41) допускает РНМ критерий для описанных в этом примере H_0 и H_1. Этот РНМ критерий основан на НКО (22.42), зависящей от $x_{(1)}$ и $\bar{x}$.

Ранее было показано (см. **17.36**, пример 17.19 и упражнение 17.9), что наименьшее наблюдение $x_{(1)}$ достаточно для θ при известном σ и что $\bar{x}$ достаточно для σ при известном θ. Пара статистик $x_{(1)}$ и $\bar{x}$ совместно достаточна, однако одномерной достаточной статистики для θ и σ не существует.

22.23 С другой стороны, следующий пример показывает, что может существовать одномерная статистика, в то время как одностороннего РНМ критерия не существует, даже когда дело касается одномерного параметра.

Пример 22.11

Рассмотрим n случайных величин $x_1, \ldots, x_n$, имеющих совместно многомерное нормальное распределение. Пусть

$$\mathbf{M}(x_1) = n\theta, \qquad \theta > 0,$$
$$\mathbf{M}(x_r) = 0, \qquad r > 1,$$

а матрица рассеяния имеет вид

$$\boldsymbol{V} = \begin{pmatrix} n-1+\theta^2 & -1 & \cdots & -1 \\ -1 & 1 & & \\ \vdots & & \ddots & 0 \\ \vdots & & 0 & \ddots \\ -1 & & & 1 \end{pmatrix}. \tag{22.46}$$

Нетрудно найти, что определитель этой матрицы равен

$$|\boldsymbol{V}| = \theta^2,$$

а обратная матрица равна

$$\boldsymbol{V}^{-1}=\frac{1}{\theta^2}\begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1+\theta^2 & & \\ \cdot & & \cdot & 1 \\ \cdot & 1 & \cdot & \\ 1 & & & 1+\theta^2 \end{pmatrix}. \tag{22.47}$$

Таким образом, в силу **15.3** получаем совместное распределение

$$dF=\frac{1}{\theta\,(2\pi)^{n/2}}\exp\left\{-\frac{1}{2}\left[\frac{n^2}{\theta^2}\,(\bar{x}-\theta)^2+\sum_{i=2}^{n}x_i^2\right]\right\}dx_1\ \dots\ dx_n. \tag{22.48}$$

Рассмотрим теперь задачу проверки гипотезы H_0: $\theta=\theta_0>0$ против H_1: $\theta=\theta_1>0$ на основе одного наблюдения. Согласно (22.6) НКО задается неравенством

$$\frac{L\,(x\,|\,\theta_0)}{L\,(x\,|\,\theta_1)}=\left(\frac{\theta_1}{\theta_0}\right)\exp\left\{-\frac{n^2}{2}\left[\frac{(\bar{x}-\theta_0)^2}{\theta_0^2}-\frac{(\bar{x}-\theta_1)^2}{\theta_1^2}\right]\right\}\leqslant k_\alpha,$$

которое равносильно неравенству

$$\frac{(\bar{x}-\theta_1)^2}{\theta_1^2}-\frac{(\bar{x}-\theta_0)^2}{\theta_0^2}\leqslant\frac{2}{n^2}\log\frac{k_\alpha\theta_0}{\theta_1}$$

или

$$\bar{x}^2\left(\theta_0^2-\theta_1^2\right)-2\bar{x}\theta_0\theta_1\,(\theta_0-\theta_1)\leqslant\frac{2\theta_0^2\theta_1^2}{n^2}\log\frac{k_\alpha\theta_0}{\theta_1}.$$

Если $\theta_0>\theta_1$, то мы получаем

$$\bar{x}^2\,(\theta_0+\theta_1)-2\bar{x}\theta_0\theta_1\leqslant a_\alpha, \tag{22.49}$$

что дает

$$b_\alpha\leqslant\bar{x}\leqslant c_\alpha. \tag{22.50}$$

В случае $\theta_0<\theta_1$ для НКО получаем

$$\bar{x}^2\,(\theta_0+\theta_1)-2\bar{x}\theta_0\theta_1\geqslant d_\alpha, \tag{22.51}$$

т. е.

$$\bar{x}\leqslant e_\alpha\quad\text{или}\quad\bar{x}\geqslant f_\alpha. \tag{22.52}$$

Как в (22.50), так и в (22.52) пределы, внутри или вне которых должно находиться $\bar{x}$, зависят от точного значения θ_1. Это обстоятельство, вызванное тем, что θ_1 входит в коэффициент при $\bar{x}^2$ в (22.49) и (22.51), означает, что даже для класса односторонних альтернатив не существует НКО, а следовательно, и РНМ критерия.

Легко проверить, что $\bar{x}$ является одномерной достаточной статистикой для θ. Этот факт завершает доказательство того, что одномерная достаточность не влечет существования РНМ критерия.

Функция мощности

22.24 Рассмотрение задачи проверки простой гипотезы H_0 против сложной гипотезы H_1 приводит к обобщению понятия мощности критерия, определенного в (22.2). Согласно этому определению мощность является явной функцией от H_1. Если, как это обычно бывает, H_1 образуется в результате изменения значений набора параметров θ, то мощность критерия для проверки гипотезы H_0: $\theta = \theta_0$, против простой альтернативы H_1: $\theta = \theta_1 \neq \theta_0$, будет зависеть от значения θ_1. Так, в примере 22.3 было показано, что мощность выборочного среднего $\bar{x}$ как критерия для проверки гипотезы о том, что среднее μ нормального распределения равно μ_0, против альтернативного значения $\mu_1 > \mu_0$ дается выражением (22.18), являющимся монотонно возрастающей функцией от μ_1. (22.18) называется *функцией мощности* критерия $\bar{x}$ для проверки H_0 против класса альтернатив H_1: $\mu > \mu_0$. Последняя запись указывает на то, что H_1 — сложная гипотеза, в противоположность записи H_1: $\mu = \mu_1 > \mu_0$, используемой для обозначения простой гипотезы.

Функция мощности редко находится так просто, как в примере 22.3, поскольку, даже если выборочное распределение точно известно как для H_0, так и для класса альтернатив H_1 (а чаще известны только приближения, особенно для H_1), остается еще задача вычисления (22.2) для каждого значения θ из H_1, которая обычно решается только с помощью численных методов. Лишь в редких случаях функция мощности находится с помощью табулированных интегралов, как это имело место в (22.18). В асимптотическом случае, однако, на помощь приходит центральная предельная теорема: распределения многих статистик критериев с возрастанием объема выборки стремятся к нормальному как при гипотезе H_0, так и при H_1, и тогда асимптотическая функция мощности, как мы увидим позже, будет иметь вид (22.18).

Пример 22.12

Общий вид функции мощности (22.18) в примере 22.3 представляет собой просто функцию нормального распределения. Она возрастает от значения

$$G(-d_\alpha) = \alpha$$

при $\mu = \mu_0$ (в соответствии с требованием размера критерия) до значения

$$G(0) = 0{,}5$$

при $\mu = \mu_0 + d_\alpha/n^{1/2}$ с одновременным возрастанием производной G'; при дальнейшем изменении μ G' убывает и стремится к нулю, а G возрастает, приближаясь к 1.

22.25. Зная функцию мощности критерия, мы можем поставить вопрос о нахождении объема выборки, необходимого для проверки гипотезы H_0 с заданным размером и мощностью. Эта процедура иллюстрируется следующим примером.

Пример 22.13

Сколько наблюдений надо взять в примере 22.3, чтобы можно было проверить гипотезу H_0: $\mu = 3$ с $\alpha = 0{,}05$ (т. е. $d_\alpha = 1{,}6449$) и мощностью, не меньшей 0,75, против альтернативы $\mu \geqslant 3{,}5$? Другими словами, как велико должно быть n, чтобы обеспечить вероятность ошибки первого рода 0,05 и вероятность ошибки второго рода не более 0,25 для $\mu \geqslant 3{,}5$?

Согласно (22.18) нам нужно найти такое n, чтобы было выполнено равенство

$$G\{n^{1/2}(3{,}5 - 3) - 1{,}6449\} = 0{,}75 \qquad (22.53)$$

(очевидно, что для $\mu > 3{,}5$ мощность будет больше). Из таблиц нормального распределения находим

$$G\{0{,}6745\} = 0{,}75 \qquad (22.54)$$

и, следовательно,

$$0{,}5n^{1/2} - 1{,}6449 = 0{,}6745,$$

откуда

$$n = (4{,}6388)^2 \approx 21{,}5,$$

так что объем выборки $n = 22$ будет достаточным для того, чтобы критерий обладал требуемой мощностью.

Односторонние и двусторонние критерии

22.26. В **22.18** было показано, что в общем случае не существует РНМ критерия, если проверяется параметрическая гипотеза H_0: $\theta = \theta_0$ против двусторонней альтернативной гипотезы, т. е. гипотезы, при которой $\theta - \theta_0$ меняет знак. Однако такая альтернативная гипотеза оказывается подходящей во многих ситуациях, например, когда нет никаких априорных сведений относительно значений θ, которые могут иметь место. В таких случаях кажется соблазнительным по-прежнему использовать статистику, приводящую к РНМ критерию, против односторонних альтернатив ($\theta > \theta_0$ или $\theta < \theta_0$), но модифицировать при этом критическую область в распределении статистики, принимая во внимание как НКО для $\theta > \theta_0$, так и НКО для $\theta < \theta_0$.

22.27. Так, в примере 22.2 и **22.17** было показано, что при проверке гипотезы H_0: $\mu = \mu_0$ о среднем μ нормального распределения среднее $\bar{x}$ дает РНМ критерий против $\mu_1 < \mu_0$ с общей для всех альтернатив НКО $\bar{x} \leqslant a_\alpha$ и РНМ критерий про-

тив $\mu_1 > \mu_0$ с общей для всех альтернатив НКО $\bar{x} \geqslant b_\alpha$. Предположим теперь, что для двусторонней альтернативы H_1: $\mu \neq \mu_0$ построена компромиссная критическая область

$$\bar{x} \leqslant a_{\alpha/2}, \qquad \bar{x} \geqslant b_{\alpha/2}. \tag{22.55}$$

Другими словами, мы объединили две односторонние критические области размера $\alpha/2$, так что размер всей критической области остался равным α.

Критическая область (22.55) всегда будет менее мощной, чем какая-то одна из односторонних НКО; однако она всегда будет более мощной, чем другая односторонняя НКО. Действительно, мощность комбинированной критической области равна (см. пример 22.3)

$$G\{n^{1/2}(\mu - \mu_0) - d_{\alpha/2}\} + G\{n^{1/2}(\mu_0 - \mu) - d_{\alpha/2}\}, \tag{22.56}$$

что представляет собой четную функцию от $(\mu - \mu_0)$ с минимумом в точке $\mu = \mu_0$. Она, следовательно, всегда заключена между $G\{n^{1/2}(\mu - \mu_0) - d_\alpha\}$ и $G\{n^{1/2}(\mu_0 - \mu) - d_\alpha\}$, являющимися функциями мощности односторонних НКО (за исключением

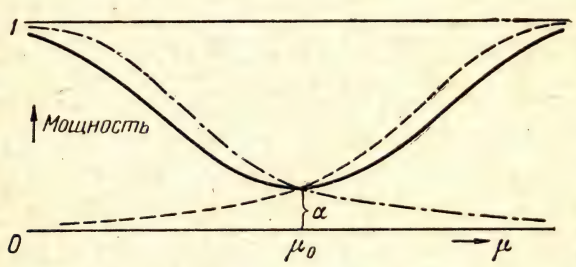


Рис. 22.3. Функции мощности трех критериев, основанных на $\bar{x}$. —— критическая область, содержащая равные «хвосты». ——— критическая область, использующая правый «хвост». —·—·— критическая область, использующая левый «хвост».

точки $\mu = \mu_0$, когда все три выражения равны). На рис. 22.3 для фиксированных значений n и α приведен соответствующий график.

22.28 Далее мы увидим, что такому способу образования критической области из областей, соответствующих «хвостам» распределения статистики критерия, может быть дано другое, менее интуитивное обоснование. Пока же эту процедуру следует рассматривать как естественный способ гарантировать себя от чрезмерных потерь мощности, которые, как показывает рис. 22.3, могут произойти при использовании любой из односторонних критических областей.

Выбор размера критерия

22.29 До сих пор в нашем изложении мы предполагали, что размер критерия α каким-то образом выбран и фиксирован. Все наши результаты сохраняют силу при любом выборе α. Теперь мы рассмотрим вопрос о том, как определять α.

Во-первых, естественно потребовать, чтобы α было «мало», в соответствии с некоторым приемлемым критерием, и, действительно, обычно используют одно из общепринятых значений α, таких, как 0,05, 0,01 или 0,001. Однако не следует заходить слишком далеко в этом направлении. Мы можем задать только две из величин n, α и β даже при проверке простой гипотезы H_0 против простой альтернативы H_1. Если n фиксировано, то, вообще говоря, мы можем уменьшить вероятность ошибки первого рода α, только увеличивая вероятность ошибки второго рода β. Другими словами, уменьшение размера критерия вызывает уменьшение мощности.

Выражение (22.18) для мощности НКО одностороннего критерия для среднего значения нормального распределения в примере 22.3 хорошо иллюстрирует этот факт. Мы видим, что при $\alpha \to 0$ в силу (22.16) $d_\alpha \to \infty$ и, следовательно, мощность (22.18) также стремится к нулю.

Таким образом, при фиксированном объеме выборки мы должны соизмерять размер и мощность критерия. Если практический риск, связанный с ошибкой первого рода, велик, в то время как риск, связанный с ошибкой второго рода, мал, то при фиксированном n следует уменьшить α за счет увеличения β.

Если, однако, объем выборки находится в нашем распоряжении, то мы можем, как в примере 22.13, взять такое достаточно большое n, чтобы уменьшить α и β до заданных уровней. Эти уровни еще надо определить, однако до тех пор, пока мы не имеем дополнительной информации в форме стоимостей (в денежных или других единицах) двух типов ошибок и стоимостей проведения наблюдений, мы не сможем получить «оптимальной» комбинации α, β и n для данной задачи. Можно лишь заметить, что независимо от способа определения α полученные в этой главе результаты остаются в силе.

22.30 Вопрос, рассмотренный в **22.29**, связан с другим, который кладется иногда в основу критики теории проверки гипотез.

Предположим, что мы применяем критерий с фиксированным α, а n очень велико. Любой разумный критерий будет иметь мощность, близкую к единице, обнаруживая любые отклонения от проверяемой гипотезы. Приведем теперь довод, сформулированный Берксоном (1938): «Никто в действитель-

ности не считает, что какая-либо гипотеза выполняется точно: мы просто строим абстрактную модель реальных событий, которая в какой-то мере обязательно отклоняется от истины. Однако, как мы видим, огромная выборка почти наверняка (т. е. с вероятностью, стремящейся к 1 при неограниченном возрастании n) отвергает гипотезу при любом заданном размере α. Зачем тогда вообще пытаться проверять гипотезу при малых выборках, которые приводят к еще менее надежному заключению, чем большие выборки?»

Этот парадокс в действительности связан с двумя моментами. Прежде всего, если n фиксировано и нас интересует не точное, а лишь приближенное выполнение проверяемой гипотезы, то это обстоятельство можно было бы учесть, сделав нашу альтернативную гипотезу достаточно удаленной от проверяемой гипотезы, чтобы отличие между ними представляло практический интерес. Само по себе это увеличило бы мощность критерия. Но если мы не желаем отвергать проверяемую гипотезу на основании небольших отклонений от нее, то следует предпочесть критерий с низкой мощностью против малых отклонений, а это означает малое α при соответственно высоком β, т. е. малой мощности.

Однако основным моментом парадокса является рассуждение, связанное с увеличением объема выборки. Проверяемая гипотеза будет отвергнута с вероятностью, близкой к 1, только если мы будем сохранять α фиксированным. Нет никаких причин, чтобы поступать таким образом: мы можем определить α любым способом, каким мы пожелаем, и в свете рассуждений пункта **22.29** разумно использовать выигрыш в чувствительности, получающийся при увеличении n, для уменьшения как β, так и α. Лишь привычка фиксировать α на определенном принятом уровне приводит к парадоксу. Если допустить уменьшение α при возрастании n, то уже нельзя быть уверенным, что очень малые отклонения от H_0 приведут к тому, что H_0 будет отвергнута: теперь это будет зависеть от того, с какой скоростью убывает α.

22.31 Парадокс, обратный парадоксу пункта **22.30**, может возникнуть при малых выборках. Так же как при больших n негибкое использование общепринятых значений α может привести к слишком высокой мощности, так и при малых фиксированных n использование таких значений α может привести к очень низкой мощности, возможно, слишком низкой. Положение может быть исправлено с помощью увеличения α, которое повлечет уменьшение β.

Статистик всегда должен убедиться в том, что в условиях его задачи он не слишком увеличил чувствительность в одном направлении за счет другого.

Пример 22.14

Э. Пирсон (обсуждение работы Линдли (1953а)) вычислил несколько значений функции мощности (22.56) двустороннего критерия для нормального среднего, которые мы воспроизводим для иллюстрации наших рассуждений.

Таблица 22.1

Функция мощности, вычисленная по формуле (22.56).
Значения в первой строке таблицы дают размеры критериев

Значение $\lvert\mu-\mu_0\rvert$	Объем выборки (n)						
	10			20	100		
0	0,050	0,072	0,111	0,050	0,050	0,019	0,0056
0,1				0,073	0,170	0,088	0,038
0,2	0,097	0,129	0,181	0,145	0,516	0,362	0,221
0,3				0,269	0,851	0,741	0,592
0,4	0,244	0,298	0,373	0,432	0,979	0,950	0,891
0,5				0,609	0,999	0,996	0,987
0,6	0,475	0,539	0,619	0,765			

Из таблицы видно, что при объеме выборки 100 уменьшение α от 0,050 до 0,019 и 0,0056 вызывает последовательное уменьшение мощности критерия для каждого значения $|\mu-\mu_0|$. Действительно, для $\alpha=0{,}0056$ и $|\mu-\mu_0|=0{,}1$ мощность падает ниже значения, соответствующего $n=20$ и $\alpha=0{,}05$. Наоборот, при уменьшении объема выборки от 20 до 10 увеличение α до 0,072 и 0,111 увеличивает мощность, хотя только в случае $\alpha=0{,}111$ и $|\mu-\mu_0|=0{,}2$ она превосходит мощность при $n=20$ и $\alpha=0{,}05$.

УПРАЖНЕНИЯ

22.1 Непосредственно используя (22.6), показать, что НКО для проверки простой гипотезы H_0: $\mu=\mu_0$ о среднем пуассоновского распределения против простой альтернативы H_1: $\mu=\mu_1$ имеет вид

$$\bar{x}\leqslant a_\alpha, \quad \text{если} \quad \mu_0>\mu_1,$$
$$\bar{x}\geqslant b_\alpha, \quad \text{если} \quad \mu_0<\mu_1,$$

где $\bar{x}$ — выборочное среднее, а a_α и b_α — константы.

22.2 Показать аналогичным образом, что НКО для параметра π биномиального распределения имеет вид

$$x\leqslant a_\alpha, \quad \text{если} \quad \pi_0>\pi_1,$$
$$x\geqslant b_\alpha, \quad \text{если} \quad \pi_0<\pi_1,$$

где x — наблюденное число «успехов» в выборке.

22.3 Показать, что в случае нормального распределения с нулевым средним и дисперсией σ^2 НКО для H_0: $\sigma = \sigma_0$ против альтернативы H_1: $\sigma = \sigma_1$ имеет вид

$$\sum_{i=1}^{n} x_i^2 \leqslant a_\alpha, \quad \text{если} \quad \sigma_0 > \sigma_1,$$

$$\sum_{i=1}^{n} x_i^2 \geqslant b_\alpha, \quad \text{если} \quad \sigma_0 < \sigma_1.$$

Показать, что мощность НКО для $\sigma_0 > \sigma_1$ равна $F\left\{\frac{\sigma_0^2}{\sigma_1^2}\chi^2_{\alpha,n}\right\}$, где $\chi^2_{\alpha,n}$ — нижняя 100α-процентная точка, а F — функция распределения χ^2 с n степенями свободы.

22.4 Показать, что в примере 22.4 мощность критерия равна 0,648 при $k_\alpha = 1$ и 0,352 при $k_\alpha = 0{,}5$. Построить графики двух соответствующих распределений Коши для иллюстрации мощности и размера каждого из критериев.

22.5 Показать, что мощность в упражнении 22.3 является монотонно возрастающей функцией от σ_0^2/σ_1^2. Проверить, используя таблицы распределения χ^2, что эта мощность является, кроме того, монотонно возрастающей функцией от n.

22.6 Показать что (22.21) выполняется для достаточных статистик, на которых основаны НКО в примерах 22.1—22.3.

22.7 Показать, что в условиях пункта **22.15** более эффективная оценка всегда дает более мощный критерий, если мощность соответствующего ей критерия превосходит 0,5.

(Сандрум, 1954.)

22.8 Показать, что для проверки гипотезы H_0: $\mu = \mu_0$ по выборке из распределения $dF = dx$, $\mu \leqslant x \leqslant \mu + 1$, имеется пара РНМ односторонних критериев и, следовательно, не существует РНМ критерия для всех альтернатив.

22.9 Показать, что статистика $\bar{x}$ в примере 22.11 нормально распределена со средним θ и дисперсией θ^2/n^2 и достаточна для параметра θ.

22.10 Показать, что для распределения в примере 22.10 не выполняется условие (2) пункта **22.21**.

22.11 Пусть σ в примере 22.9 — любая положительная возрастающая функция от θ. Показать, что для проверки гипотезы H_0: $\theta = \theta_0$ против H_1: $\theta = \theta_1 < \theta_0$ по прежнему имеется общая НКО.

(Нейман, Пирсон, 1936а.)

22.12 Обобщая результаты пункта **22.27**, выписать выражение для функции мощности критерия, основанного на распределении $\bar{x}$ и имеющего критическую область вида $\bar{x} \leqslant a_{\alpha_1}$, $\bar{x} \geqslant b_{\alpha_2}$, где $\alpha_1 + \alpha_2 = \alpha$ (α_1 и α_2 не обязательно равны). Показать, что функция мощности любого такого критерия целиком лежит между функциями мощности, соответствующими $\alpha_1 = 0$ и $\alpha_2 = 0$ и изображенными на рис. 22.3.

22.13 В связи с результатами пункта **22.14** показать, записывая функцию правдоподобия в виде

$$L(x \mid \theta) = L(x \mid \theta_1)\left[\frac{L(x \mid \theta_0)}{L(x \mid \theta_1)}\right]^{(\theta-\theta_1)/(\theta_0-\theta_1)},$$

что отношение правдоподобия (для проверки простой гипотезы H_0: $\theta = \theta_0$ против простой альтернативы H_1: $\theta = \theta_1$) является достаточной статистикой для параметра θ при каждой из гипотез.

(Питмэн, 1957)

ГЛАВА 23

ПРОВЕРКА ГИПОТЕЗ: СЛОЖНЫЕ ГИПОТЕЗЫ

23.1 В главе 22 мы видели, что если проверяемая гипотеза проста (т. е. определяет распределение полностью), то *всегда* имеется НКО, дающая наиболее мощный критерий против простой альтернативной гипотезы, а в случае альтернативы, представляющей собой класс простых параметрических гипотез, *может* существовать РНМ критерий, однако, вообще говоря, РНМ критерия не существует, если значения параметра, порождающие альтернативную гипотезу, могут лежать по обе стороны от значения, отвечающего проверяемой гипотезе.

Если же проверяется сложная гипотеза, оставляющая неопределенным значение по крайней мере одного параметра, то следует ожидать, что РНМ критерии будут существовать еще реже, чем в случае проверки простой гипотезы. Однако мы увидим, что определенный прогресс может быть достигнут за счет сужения тем или иным способом класса рассматриваемых критериев.

Сложные гипотезы

23.2 Прежде всего сформулируем точно нашу задачу. Предполагается, что n наблюдений имеют распределение, зависящее от l ($\leqslant n$) параметров, которое мы, как и прежде, будем обозначать

$$L(x|\,\theta_1, \ldots, \theta_l).$$

Проверяемая гипотеза имеет вид

$$H_0\colon\ \theta_1 = \theta_{10};\quad \theta_2 = \theta_{20};\ \ldots;\ \theta_k = \theta_{k0}, \tag{23.1}$$

где $k \leqslant l$, а второй индекс 0 указывает на значение, задаваемое гипотезой. То, что гипотезой задаются значения первых k из l параметров, не ограничивает общности. Говорят, что гипотеза H_0, определенная в (23.1), накладывает k ограничений или имеет $l - k$ степеней свободы, хотя при употреблении этого последнего (более старого) термина следует быть осмотрительным, поскольку мы уже используем термин «степени свободы» в другом смысле.

Гипотеза вида

$$H_0\colon\ \theta_1 = \theta_2;\quad \theta_3 = \theta_4;\ \ldots,$$

не определяющая значений параметров, равенство которых проверяется, может быть преобразована к виду (23.1) с помощью введения параметров $\theta_1 - \theta_2$, $\theta_3 - \theta_4$ и т. д. и проверки гипотезы о равенстве нулю новых параметров. Таким образом, сложная гипотеза (23.1) обладает большей общностью, чем может показаться на первый взгляд.

Чтобы упростить обозначения, мы будем писать $L(x\,|\,\theta_r, \theta_s)$ и

$$H_0\colon\ \theta_r = \theta_{r0}, \tag{23.2}$$

где параметры θ_r, θ_s могут быть многомерными, причем θ_s обозначает «мешающий параметр», значение которого проверяемой гипотезой не определяется.

Оптимальное свойство достаточных статистик

23.3 Сейчас уместно доказать свойство оптимальности достаточных статистик, аналогичное свойству, полученному в **17.35**. Там было показано, что если t_1 — несмещенная оценка параметра θ и t — достаточная статистика для θ, то статистика $\mathbf{M}(t_1\,|\,t)$ является несмещенной оценкой для θ с дисперсией, не превосходящей дисперсии t_1. Теперь мы докажем следующий результат, принадлежащий Леману: если w — критическая область для гипотезы H_0 о параметре θ распределения $L(x\,|\,\theta)$ против некоторой альтернативы H_1 и t — достаточная статистика для θ как при H_0, так и при H_1, то существует критерий такого же размера, основанный на некоторой функции от t и имеющей мощность, равную мощности w.

Определим функцию *)

$$c(w) = \begin{cases} 1, & \text{если выборочная точка принадлежит } w, \\ 0 & \text{в противном случае.} \end{cases} \tag{23.3}$$

Тогда интеграл

$$\int c(w)\, L(x\,|\,\theta)\, dx = \mathbf{M}\{c(w)\} \tag{23.4}$$

даст вероятность того, что выборочная точка попадет в w, и, следовательно, равен размеру критерия α, когда верна гипотеза H_0, и равен мощности критерия, когда верна H_1. Используя

*) $c(w)$ известна в теории меры как характеристическая функция множества точек w. Мы избегаем этого термина, поскольку имеется опасность смешения с понятием «характеристическая функция», обозначающим преобразование Фурье функции распределения (глава 4).

свойство факторизации функции правдоподобия (17.68), выполняющееся при наличии достаточной статистики, можно записать (23.4) в виде

$$\mathbf{M}\{c(w)\} = \int c(w)\, h(x|t)\, g(t|\theta)\, dx = \mathbf{M}\{\mathbf{M}(c(w)|t)\}, \quad (23.5)$$

где математическое ожидание вне фигурных скобок берется по распределению статистики t. Таким образом, функция от t $\mathbf{M}(c(w)|t)$ не зависит от θ, поскольку t — достаточная статистика, и имеет математическое ожидание, равное математическому ожиданию $c(w)$. Следовательно, существует основанный на достаточной статистике t критерий, имеющий размер и мощность исходной области w *). Учитывая этот факт, при рассмотрении любой задачи о проверке гипотез можно ограничиться функциями от достаточных статистик, не теряя при этом мощности.

Это достаточно общий результат, охватывающий, в частности, случай проверки простой гипотезы H_0, изученный в главе 22.

Размер критерия для сложных гипотез: подобные области

23.4 Поскольку сложная гипотеза оставляет неопределенными значения некоторых параметров, то немедленно возникает новая проблема, связанная с тем, что размер критерия для гипотезы H_0 будет, очевидно, в общем случае зависеть от значений мешающего параметра θ_s.

Если мы хотим, чтобы ошибка первого рода была не выше некоторого заданного уровня, то мы должны найти критическую область, размер которой не превышал бы этого уровня при любом возможном значении θ_s. Таким образом, требуется, чтобы было выполнено неравенство

$$\alpha(\theta_s) \leqslant \alpha. \quad (23.6)$$

Если для всех θ_s имеет место точное равенство

$$\alpha(\theta_s) = \alpha, \quad (23.7)$$

то мы получаем (критическую) область, *подобную выборочному пространству* *) относительно θ_s, или, короче, *подобную* (критическую) область. Критерий, основанный на подобной критической области, называется подобным критерием.

---

*) Этот критерий состоит в том, что гипотеза отвергается с вероятностью $\mathbf{M}(c(w)|t)$, зависящей от наблюденного значения t, и является, вообще говоря, рандомизованным критерием с критической функцией $\mathbf{M}(c(w)|t)$ (см. примечание на стр. 226). (*Прим. ред.*)

**) Возникновение этого термина связано с тем, что все выборочное пространство, тривиальным образом, является подобной областью с $\alpha = 1$.

23.5 Совсем не очевидно, существуют ли в общем случае подобные области вообще. Однако, как указал Феллер (1938), они существуют в случае выборки, состоящей из n одинаково распределенных независимых наблюдений непрерывной варианты x. Действительно, независимо от вида распределения и значений его параметров всегда имеет место соотношение

$$\mathbf{P}\{x_1 < x_2 < x_3 < \ldots < x_n\} = 1/n! \qquad (23.8)$$

(см. **11.4**), поскольку все $n!$ перестановок величин x_i одинаково вероятны. Таким образом, для любого α, кратного $1/n!$, имеются подобные области, основанные на $n!$ гиперсимплексах в выборочном пространстве, полученных с помощью перестановок n индексов в (23.8).

23.6 Если ограничиться областями, определяемыми симметричными функциями от наблюдений (так что подобные области, основанные на (23.8), будут тогда исключены из рассмотрения), то нетрудно увидеть, что такие подобные области могут существовать для одних объемов выборок и не существовать для других. Например, для выборки объема n из нормального распределения

$$dF(x) = (2\pi)^{-1/2} \exp\left\{-\tfrac{1}{2}(x-\theta)^2\right\} dx$$

не существует подобной области относительно θ при $n = 1$. Однако при $n \geqslant 2$ из того факта, что $ns^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2$ имеет распределение χ^2 с $n-1$ степенями свободы при любом θ, следует, что подобные области любого размера могут быть найдены с помощью распределения статистики s^2. Это происходит потому, что $\bar{x}$ служит одномерной достаточной статистикой для θ, и, чтобы найти подобную область, надо, согласно упражнению 23.3, найти статистику, некоррелированную с

$$\frac{\partial \log L}{\partial \theta} = n(\bar{x} - \theta).$$

Это невозможно сделать при $n = 1$, поскольку $\bar{x} = x$ совпадает в этом случае со всей выборкой. Однако при $n \geqslant 2$ сумма $\sum (x - \bar{x})^2$ распределена независимо от $\bar{x}$ и дает, следовательно, подобные области. Аналогичная ситуация наблюдается в упражнении 23.1, где имеется пара достаточных статистик для двух параметров и требуется по меньшей мере три наблюдения, чтобы иметь статистику, независимую от обеих достаточных статистик.

23.7 Даже при большом n симметричные подобные области могут не существовать, если добавление каждого наблюдения связано с добавлением нового параметра, как это имеет место в следующем примере, принадлежащем Феллеру (1938).

Пример 23.1

Рассмотрим выборку из n наблюдений, в которой i-е наблюдение имеет распределение

$$dF(x_i) = (2\pi)^{-1/2} \exp\left\{-\frac{1}{2}(x_i - \theta_i)^2\right\} dx_i,$$

так что

$$L(x|\theta) = (2\pi)^{-n/2} \exp\left\{-\frac{1}{2}\sum_i (x_i - \theta_i)^2\right\}.$$

Для подобной области w размера α тождественно по θ выполняется равенство

$$\int_w L(x|\theta)\, dx = \alpha.$$

Используя (23.3), можно переписать это условие в виде

$$\int_W L(x|\theta) \frac{c(w)}{\alpha}\, dx = 1, \tag{23.9}$$

где W — все выборочное пространство. Дифференцируя (23.9) по θ_i, находим

$$\int_W L(x|\theta) \frac{c(w)}{\alpha}(x_i - \theta_i)\, dx = 0. \tag{23.10}$$

Второе дифференцирование по θ_i дает

$$\int_W L(x|\theta) \frac{c(w)}{\alpha}\{(x_i - \theta_i)^2 - 1\}\, dx = 0. \tag{23.11}$$

Далее, из определения $c(w)$ следует, что

$$g(x|\theta) = L(x|\theta) \frac{c(w)}{\alpha} \tag{23.12}$$

является некоторой (совместной) плотностью распределения. Соотношения (23.10) и (23.11) выражают тот факт, что маргинальные распределения x_i в $g(x|\theta)$ имеют математические ожидания и дисперсии

$$\mathsf{M}(x_i) = \theta_i, \quad \mathsf{D}x_i = 1$$

такие же, как сами величины x_i.

Изучение распределения $g(x|\theta)$ приводит к выводу, что, производя дальнейшие дифференцирования, мы бы нашли, что *все* моменты $g(x|\theta)$ совпадают с моментами распределения $L(x|\theta)$ (которое однозначно определяется своими моментами). Таким образом, из (23.10) следует тождество $c(w)/\alpha = 1$. Но поскольку функция $c(w)$ может быть равна только 0 или 1,

мы приходим к заключению, что подобные области могут существовать лишь для двух тривиальных значений α, равных 0 или 1. Трудность в данном случае заключается в том, что для образования достаточного набора статистик для n параметров требуются все n наблюдений, так что найти независимую от них статистику невозможно.

23.8 И все-таки для многих задач проверки сложных гипотез подобные области существуют при любом размере критерия α и любом объеме выборки n. Мы рассмотрим теперь методы нахождения таких областей.

Пусть t — достаточная статистика для параметра θ_s, значение которого не определяется гипотезой H_0, и пусть имеется критическая область w такая, что для всех значений t при выполнении гипотезы H_0 справедливо соотношение

$$\mathbf{M}\{c(w) \mid t\} = \alpha. \tag{23.13}$$

Беря математическое ожидание относительно t, мы, аналогично (23.5), получаем равенство

$$\mathbf{M}\{c(w)\} = \mathbf{M}\{\mathbf{M}(c(w) \mid t)\} = \alpha, \tag{23.14}$$

из которого следует, что исходная критическая область w является подобной областью размера α (Нейман (1937b) и Бартлетт (1937)). Таким образом, w содержит долю α вероятностной меры каждого контура, соответствующего постоянному значению t.

Следует заметить, что в данном случае требуется, чтобы статистика t была достаточной только для мешающего параметра θ_s и только когда верна H_0. В **23.3** требования были более жесткими.

Из наших рассуждений вытекает, что (23.13) служит достаточным условием того, чтобы w была подобной. В **23.19** мы покажем, что (23.13) при одном дополнительном требовании будет необходимым и достаточным условием. Чтобы сформулировать это требование, мы введем, следуя Леману и Шеффе (1950), понятие полноты параметрического семейства распределений. Это понятие позволит нам также продолжить изучение достаточных статистик, начатое в главе 17.

Полные параметрические семейства и полные статистики

23.9 Рассмотрим параметрическое семейство (одномерных или многомерных) распределений $f(x \mid \theta)$, зависящих от вектора параметров θ. Пусть $h(x)$ — любая статистика, не зависящая от θ. Если выполнение соотношения

$$\mathbf{M}\{h(x)\} = \int h(x) f(x \mid \theta)\, dx = 0 \tag{23.15}$$

при всех θ влечет равенство

$$h(x) = 0, \tag{23.16}$$

выполняющееся тождественно с точностью до множества меры нуль, то семейство $f(x|\theta)$ называется *полным*. Если (23.15) влечет (23.16) только для ограниченных $h(x)$, то семейство $f(x|\theta)$ называется *ограниченно полным*.

В статистических приложениях понятия полноты интересующее нас семейство распределений $g(t|\theta)$ часто представляет собой выборочное распределение некоторой статистики (возможно, векторной) при различных значениях θ. В этом случае статистика t называется полной (или ограниченно полной), если выполнение соотношения $\mathbf{M}\{h(t)\} = 0$ при всех θ влечет тождество $h(t) = 0$ для всех (или всех ограниченных) функций $h(t)$. Другими словами, полнота статистики t определяется как полнота ее распределения.

Из полноты статистики t, очевидно, сразу следует, что только одна функция от этой статистики может иметь заданное математическое ожидание. Поэтому, если какая-то функция от t служит несмещенной оценкой некоторой функции от θ, то никакая другая функция от t не будет обладать этим свойством. Таким образом, полнота обеспечивает единственность несмещенной оценки.

Гхош и Сингх (1966), используя одну теорему Винера, показали, что если θ — параметр расположения, т. е. $f = f(x - \theta)$, то ограниченная полнота эквивалентна необращению в нуль при всех t характеристической функции $\varphi(t)$. Таким образом, например, распределение Коши в примере 17.7 ограниченно полно (х. ф. этого распределения дана в примере 4.2, том 1).

Полнота достаточных статистик

23.10 В частном случае экспоненциального семейства (17.83) при $A(\theta) = \theta$ и $B(x) = x$ получаем

$$f(x|\theta) = \exp\{\theta x + C(x) + D(\theta)\}, \quad -\infty \leqslant x \leqslant \infty. \tag{23.17}$$

Если

$$\int h(x) f(x|\theta)\, dx = 0$$

при всех θ, то

$$\int [h(x) \exp\{C(x)\}] \exp(\theta x)\, dx = 0. \tag{23.18}$$

Интеграл в (23.18) есть двустороннее преобразование Лапласа *) функции, написанной под интегралом в квадратных скобках. В силу единственности преобразования существует только одна функция, имеющая в качестве преобразования нуль, а именно функция, тождественно равная нулю. Таким образом, тождественно по x

$$h(x)\exp\{C(x)\} = 0,$$

откуда следует, что тождественно по x

$$h(x) = 0,$$

т. е. семейство $f(x|\theta)$ полно.

Как показали Леман и Шеффе (1955), этот результат обобщается на многопараметрический случай: экспоненциальное семейство с k параметрами и k переменными

$$f(\boldsymbol{x}\,|\,\theta) = \exp\left\{\sum_{j=1}^{k}\theta_j x_j + C(\boldsymbol{x}) + D(\theta)\right\} \tag{23.19}$$

полно. В упражнении 17.14 мы нашли совместное распределение статистик, совместно достаточных для k параметров одномерного экспоненциального семейства (17.86). Формула (23.19) получается как частный случай этого распределения при $A_j(\theta) = \theta_j$. (Величинам nD и h из упражнения 17.14 здесь соответствуют D и $\exp(C)$.) В соответствии с **23.3** при проверке гипотез о параметрах мы можем ограничиться рассмотрением только достаточных статистик.

Пример 23.2

Рассмотрим семейство нормальных распределений

$$f(x\,|\,\theta_1, \theta_2) = (2\pi\theta_2)^{-1/2}\exp\left\{-\frac{1}{2\theta_2}(x-\theta_1)^2\right\},$$
$$-\infty \leqslant x \leqslant \infty, \quad \theta_2 > 0.$$

---

*) Двустороннее преобразование Лапласа функции $g(x)$ определяется формулой

$$\lambda(\theta) = \int_{-\infty}^{\infty}\exp(\theta x)\, g(x)\, dx.$$

Интеграл сходится в полосе $\alpha < \Re(\theta) < \beta$ комплексной плоскости. (Одна из границ α или β или обе сразу могут быть бесконечными. Полоса может также вырождаться в линию.) Между $g(x)$ и $\lambda(\theta)$ существует (с точностью до множества меры нуль) взаимно однозначное соответствие. См., например, D. V. Widder (1941), *The Laplace Transform*, Princeton U. P. Ср. также с теоремой обращения для х. ф. в **4.3.**

(а) Если θ_2 известно (скажем, равно 1), то семейство полно относительно θ_1, так как тогда мы имеем дело с частным случаем семейства (23.17), где

$$\theta = \theta_1, \quad \exp\{C(x)\} = (2\pi)^{-1/2} \exp\left(-\tfrac{1}{2}x^2\right),$$

$$D(\theta) = \exp\left(-\tfrac{1}{2}\theta_1^2\right).$$

(б) Если, с другой стороны, известно θ_1 (скажем, равно 0), то семейство не будет даже ограниченно полным относительно θ_2, так как $f(x|0, \theta_2)$ — четная функция от x, и поэтому любая нечетная функция $h(x)$ будет иметь нулевое математическое ожидание, не являясь в то же время тождественным нулем.

23.11 В **23.10** мы интересовались полнотой характеристической формы совместного распределения достаточных статистик для выборок из распределения, сосредоточенного на множестве, не зависящем от параметров. Хогг и Крэйг (1956) доказали полноту достаточной статистики для распределения, определенного на интервале, границы которого зависят от единственного параметра θ, и обладающего одномерной достаточной статистикой. В **17.40—41** было доказано, что в этом случае исходное распределение должно иметь вид *)

$$f(x \mid \theta) = q(x)/h(\theta) \tag{23.20}$$

и что

(I) если только одна граница интервала, на котором задана $f(x|\theta)$, является функцией от θ (без потери общности можно считать, что эта функция есть просто θ), то соответствующая экстремальная статистика будет достаточной;

(II) если обе границы зависят от θ и верхняя граница $b(\theta)$ является монотонно убывающей функцией от нижней границы θ, то существует одномерная достаточная статистика, имеющая вид

$$\min\{x_{(1)}, b^{-1}(x_{(n)})\}. \tag{23.21}$$

Мы рассмотрим отдельно каждый из случаев (I) и (II).

23.12 Пусть в случае (I) верхняя граница равна θ, а нижняя постоянна и равна a. Тогда статистика $x_{(n)}$ достаточна для θ, а

*) В английском оригинале числитель правой части в (23.20) обозначен, как и в **17.40**, через $g(x)$. В переводе это обозначение изменено на $q(x)$, поскольку $g(x)$ обозначает также плотность распределения достаточной статистики (см., например, (23.26)). (*Прим. ред.*)

ее распределение согласно (11.34) и (23.20) имеет вид

$$dG(x_{(n)}) = n\{F(x_{(n)})\}^{n-1} f(x_{(n)})\, dx_{(n)} = \\ = \frac{n\left\{\int\limits_a^{x_{(n)}} q(x)\, dx\right\}^{n-1} q(x_{(n)})}{\{h(\theta)\}^n} dx_{(n)}, \quad a \leqslant x_{(n)} \leqslant \theta. \qquad (23.22)$$

Предположим теперь, что для некоторой статистики $u(x_{(n)})$ выполнено соотношение

$$\int\limits_a^{\theta} u(x_{(n)})\, dG(x_{(n)}) = 0,$$

которое, используя (23.22) и опуская множитель с $h(\theta)$, можно записать в виде

$$\int\limits_a^{\theta} u(x_{(n)}) \left\{\int\limits_a^{x_{(n)}} q(x)\, dx\right\}^{n-1} q(x_{(n)})\, dx_{(n)} = 0. \qquad (23.23)$$

Дифференцируя (23.23) по θ, находим

$$u(\theta)\left\{\int\limits_a^{\theta} q(x)\, dx\right\}^{n-1} q(\theta) = 0. \qquad (23.24)$$

Поскольку интеграл в фигурных скобках равен $h(\theta)$ и $q(\theta) \neq \\ \neq 0 \neq h(\theta)$, то из (23.24) следует, что для любого θ

$$u(\theta) = 0.$$

Таким образом, функция $u(x_{(n)})$ тождественно равна нулю и распределение (23.22) статистики $x_{(n)}$ полно. Приведенное доказательство справедливо также в случае, когда от параметра зависит нижняя граница.

23.13 В случае (II) функция распределения достаточной статистики (23.21) имеет вид

$$G(t) = 1 - \mathbf{P}\{x_{(1)},\ b^{-1}(x_{(n)}) > t\} = 1 - \mathbf{P}\{x_{(1)} > t,\ x_{(n)} < b(t)\} = \\ = 1 - \left\{\int\limits_t^{b(t)} \frac{q(x)}{h(\theta)}\, dx\right\}^n. \qquad (23.25)$$

Дифференцируя ее по t, находим плотность

$$g(t) = \frac{n}{\{h(\theta)\}^n}\left\{\int\limits_t^{b(t)} q(x)\, dx\right\}^{n-1} [q(t) - q\{b(t)\}\, b'(t)], \qquad (23.26)$$

$$\theta \leqslant t \leqslant c(\theta).$$

Пусть теперь для некоторой статистики $u(t)$ выполняется соотношение

$$\int_{\theta}^{c(\theta)} u(t)\, g(t)\, dt = 0. \qquad (23.27)$$

Дифференцируя (23.27) по θ и повторяя рассуждения пункта **23.12**, получаем, что $u(\theta) = 0$ для любого θ. Таким образом, функция $u(t)$ тождественно равна 0 и распределение $g(t)$, определяемое формулой (23.26), полно.

23.14 Ниже приводится пример неполной одномерной достаточной статистики.

Пример 23.3

Рассмотрим выборку, состоящую из одного наблюдения случайной величины, имеющей прямоугольное распределение $dF = dx$, $\theta \leqslant x \leqslant \theta + 1$.

Очевидно, x — достаточная статистика. (В случае $n \geqslant 2$ одномерной достаточной статистики не существует, поскольку не выполнено условие (II) пункта 23.11.)

Для любой периодической с периодом 1 функции $h(x)$, удовлетворяющей соотношению

$$\int_{0}^{1} h(x)\, dx = 0,$$

получаем

$$\int_{\theta}^{\theta+1} h(x)\, dF = \int_{\theta}^{\theta+1} h(x)\, dx = \int_{0}^{1} h(x)\, dx = 0.$$

Отсюда следует, что рассматриваемое распределение не является даже ограниченно полным, поскольку $h(x)$ не равна тождественно нулю.

Минимальная достаточность

23.15 В **17.38** указывалось, что, когда мы рассматриваем достаточные статистики в общем случае (т. е. не ограничиваясь, как это делалось ранее в главе 17, случаем одномерной достаточности), нам приходится выбирать между различными системами достаточных статистик. В случае выборки объема n *всегда* имеется система из n достаточных статистик (а именно сами n наблюдений) для k ($\geqslant 1$) параметров распределения, из которого извлечена выборка. Иногда, хотя и не всегда, имеется

достаточная система из s ($<n$) статистик. Часто $s = k$, например, это имеет место во всех случаях достаточности, рассмотренных в примерах 17.15, 17.16, где $s = k = 1$, и в примере 17.17, где $s = k = 2$. В противоположность этому, в следующем примере рассматривается случай $s < k$ *).

Пример 23.4

Рассмотрим снова задачу примера 22.11 с единственным изменением, состоящим в том, что $\mathbf{M}(x_1) = n\mu$, а не $n\theta$, как прежде. Точно так же, как раньше, находим совместное распределение

$$dF = \frac{1}{\theta\,(2\pi)^{n/2}} \exp\left\{ -\frac{1}{2}\left[\frac{n^2}{\theta^2}(\bar{x} - \mu)^2 + \sum_{i=2}^{n} x_i^2 \right] \right\} dx_1 \ldots dx_n.$$

Ясно, что статистика $\bar{x}$ одна достаточна для параметров μ, θ.

23.16 Итак, возникает вопрос о том, каково наименьшее число s статистик, образующих достаточную систему для некоторой задачи. Леман и Шеффе (1950) определяют минимальный достаточный вектор статистик как вектор, являющийся функцией от всех других векторов, достаточных для параметров распределения **). Каким образом, однако, можно в любой конкретной задаче проверить, является ли некоторый достаточный вектор минимальным? И можно ли указать способ построения минимального достаточного вектора?

Частичный ответ на первый вопрос дает следующий результат. Если вектор $\boldsymbol{t}$ является ограниченно полной достаточной статистикой для $\boldsymbol{\theta}$, а вектор $\boldsymbol{u}$ есть минимальная достаточная статистика для $\boldsymbol{\theta}$, то $\boldsymbol{t}$ и $\boldsymbol{u}$ эквивалентны, т. е. они могут отличаться лишь на множестве меры нуль.

Доказывается это просто. Пусть w — область в выборочном пространстве, для которой

$$D = \mathbf{M}\,(c\,(w)\,|\,\boldsymbol{t}) - \mathbf{M}\,(c\,(w)\,|\,\boldsymbol{u}) \neq 0, \tag{23.28}$$

где функция $c(w)$ определена в (23.3). Беря в (23.28) математические ожидания по всему выборочному пространству, получаем

$$\mathbf{M}\,(D) = 0. \tag{23.29}$$

*) Фишер (например, 1956) называет достаточную систему статистик «достаточной», только если $s = k$. Если $s > k$, то он называет достаточную систему «исчерпывающей» (exhaustive).

**) Как показали Баранкин и Катц (1959), в практических ситуациях минимальный достаточный вектор есть достаточный вектор с минимальным числом компонент. См. также Баранкин (1960a, 1960b, 1961) и Фрэзер (1963).

Так как $\boldsymbol{u}$ — минимальная достаточная статистика, то она по определению является функций от другой достаточной статистики t. Следовательно, (23.28) можно записать в виде

$$D = h(t) \neq 0. \tag{23.30}$$

Однако поскольку D — ограниченная функция, то (23.29) и (23.30) противоречат предположению об ограниченной полноте t. Таким образом, не существует области w, для которой выполнялось бы (23.28), и, следовательно, t и $\boldsymbol{u}$ эквивалентны, т. е. t — минимальная достаточная статистика.

Обратное неверно: в то время как ограниченная полнота влечет минимальную достаточность, минимальная достаточность может существовать без ограниченной полноты. В этом отношении поучителен пример 23.10, приводимый ниже.

Как следствие доказанного результата получаем, что не может быть более одной ограниченной полной достаточной статистики. В примере 23.2(б) минимальная достаточная статистика x^2 полна, тогда как x достаточна, но неполна.

Другая постановка задачи о минимальной достаточности дана Дынкиным (1951).

23.17 В силу результатов **23.10—13**, относящихся к полноте достаточных статистик, из **23.16** следует, что все достаточные статистики, рассмотренные нами в предыдущих главах, являются, как и следовало ожидать из интуитивных соображений, минимально достаточными.

23.18 Более прямым способом нахождения минимальной достаточной статистики, чем метод пункта **23.16**, служит следующая процедура, предложенная Леманом и Шеффе (1950).

В **22.14** и **22.20** было показано, что при проверке простой гипотезы отношение правдоподобия является функцией достаточной статистики (системы достаточных статистик). Теперь мы можем, так сказать, обратить этот результат и использовать его для нахождения минимальной достаточной системы. Обозначая, как и прежде, функцию правдоподобия $L(x|\theta)$ (x и θ могут быть векторами), рассмотрим некоторое значение x_0 и найдем все допустимые значения x, для которых $L(x|\theta) \neq 0$ и отношение

$$\frac{L(x|\theta)}{L(x_0|\theta)} = k(x, x_0) \tag{23.31}$$

не зависит от θ. Любая достаточная (возможно, векторная) статистика t удовлетворяет (17.68), поэтому

$$\frac{L(x|\theta)}{L(x_0|\theta)} = \frac{g(t|\theta)}{g(t_0|\theta)} \cdot \frac{h(x)}{h(x_0)}, \tag{23.32}$$

так что если $t = t_0$, то (23.32) имеет вид (23.31). Обратно, если

(23.31) выполняется для всех θ, то достаточная статистика t постоянна и равна t_0. Это свойство может быть использовано для распознавания достаточных статистик и нахождения минимальной системы, как это сделано в следующих примерах.

Пример 23.5

В примере 17.17 было показано, что система статистик $(\bar{x}, s^2)$ совместно достаточна для параметров (μ, σ^2) нормального распределения. Для этого распределения $L(x|\theta) \neq 0$ при всех $\sigma^2 > 0$. Условие (23.31) после логарифмирования сводится к требованию, чтобы выражение

$$-\frac{1}{2\sigma^2}\left\{\left(\sum_i x_i^2 - \sum_i x_{0i}^2\right) - 2\mu n(\bar{x} - \bar{x}_0)\right\} \qquad (23.33)$$

не зависело от (μ, σ^2), т. е. чтобы выражение в фигурных скобках было равно нулю. Это будет выполнено, например, если каждое x_i равно соответствующему x_{0i}, что лишний раз доказывает совместную достаточность n наблюдений, которая, как мы отмечали, имеет место всегда.

Условие (23.33) будет также выполнено, если x_i образуют некоторую перестановку значений x_{0i}; таким образом, система *порядковых статистик* достаточна, что, впрочем, также всегда справедливо. Наконец, это условие будет выполнено, если

$$\bar{x} = \bar{x}_0, \qquad \sum_i x_i^2 = \sum_i x_{0i}^2, \qquad (23.34)$$

и ясно, что еще меньшей системы, для которой оно было бы выполнено, не существует. Таким образом, пара $(\bar{x}, \sum x^2)$ образует минимальную достаточную систему. Поскольку $ns^2 = \sum x^2 - n\bar{x}^2$, то пара $(\bar{x}, s^2)$ будет также минимальной достаточной.

Пример 23.6

К другой ситуации мы приходим в случае распределения Коши, рассмотренном в примере 17.7. Функция правдоподобия $L(x|\theta)$ для распределения Коши всюду отлична от нуля, и (23.31) соответствует требованию, чтобы отношение

$$\prod_{i=1}^{n}\{1 + (x_{0i} - \theta)^2\} \Big/ \prod_{i=1}^{n}\{1 + (x_i - \theta)^2\} \qquad (23.35)$$

не зависело от θ. Как и в предыдущем примере, система порядковых статистик достаточна, однако меньшей системы найти не удается. Действительно, поскольку (23.35) является отношением двух многочленов степени $2n$ по θ, то для незави-

симости от θ каждый из них должен иметь (с точностью до перестановки) один и тот же набор корней. Таким образом, мы не можем найти ничего лучшего, чем порядковые статистики, в качестве минимальной достаточной системы.

Полнота и подобные области

23.19 Мы заканчиваем наше длинное отступление, относящееся к полноте, и возвращаемся к изучению подобных областей, начатому в **23.8**. Мы можем теперь показать, что если при гипотезе H_0 достаточная статистика t ограниченно полна, то все подобные области размера α будут удовлетворять соотношению (23.13). Действительно, для любой такой области выполняется соотношение (23.14), которое можно записать в виде

$$\mathbf{M}\{\mathbf{M}(c(w)\,|\,t) - \alpha\} = 0. \qquad (23.36)$$

Выражение в фигурных скобках ограничено, так что если статистика t ограниченно полна, то из (23.36) следует, что выполняется тождество

$$\mathbf{M}(c(w)\,|\,t) - \alpha = 0,$$

т. е. имеет место (23.13).

Как показали Леман и Шеффе (1950), справедливо и обратное: если для всех подобных областей выполнено (23.13), то статистика t будет ограниченно полной. Таким образом, ограниченная полнота достаточной статистики эквивалентна выполнению условия (23.13) для всех подобных областей w.

Выбор наиболее мощной подобной области

23.20 Важное значение результата пункта **23.19** состоит в том, что он позволяет свести задачу нахождения наиболее мощной подобной области для сложной гипотезы к уже знакомой задаче нахождения НКО для простой гипотезы.

Согласно **23.19** из ограниченной полноты статистики t, достаточной для θ_s при H_0, следует, что для всех подобных областей w выполняется (23.13), т. е. любая подобная область содержит долю α вероятностной меры каждого контура, соответствующего постоянному значению t. Мы, следовательно, можем считать в наших рассмотрениях статистику t постоянной. Из постоянства статистики t, достаточной для θ_s, в силу (17.68) следует, что условное распределение наблюдений в выборочном пространстве не будет зависеть от θ_s. Таким образом, сложная гипотеза H_0 сводится к простой гипотезе H_0 при постоянном значении t. Если, кроме того, t достаточна для θ_s при гипотезе

H_1, то сложная гипотеза H_1 также сводится к простой при постоянном значении t (при этом мощность и размер любой критической области не будут зависеть от θ_s). Если, однако, t не является достаточной для θ_s при гипотезе H_1, то H_1 следует рассматривать как класс простых альтернатив к гипотезе H_0, точно так же, как это делалось в предыдущей главе.

Таким образом, сохраняя значение t постоянным, мы сводим нашу задачу к проверке простой гипотезы H_0, относящейся к θ_r, против простой альтернативы H_1 (или класса простых альтернатив, образующих H_1). Для нахождения НКО (или общей НКО) для проверки H_0 против H_1 можно использовать методы предыдущей главы, основанные на лемме Неймана — Пирсона (22.6). Если такая НКО имеется для каждого фиксированного значения t, то она, очевидно, будет безусловной НКО, дающей наиболее мощный подобный критерий для H_0 против H_1. Как и раньше, если этот критерий остается наиболее мощным против класса альтернативных значений θ_r, он называется РНМ подобным критерием.

Пример 23.7

Проверить H_0: $\mu = \mu_0$ против H_1: $\mu = \mu_1$ в случае нормального распределения

$$dF = \frac{1}{\sigma (2\pi)^{1/2}} \exp\left\{ -\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \right\} dx, \qquad -\infty \leqslant x \leqslant \infty.$$

H_0 и H_1 — сложные гипотезы с одной степенью свободы, так как значение σ^2 не определено. Из примеров 17.10 и 17.15 следует, что статистика $u = \sum_{i=1}^{n} (x_i - \mu_0)^2$ (вычисленная по выборке из n независимых наблюдений) достаточна для σ^2 при гипотезе H_0 и только при этой гипотезе. Согласно **23.10** u — полная статистика. Следовательно, любая подобная область для H_0 содержит долю α каждого контура, соответствующего постоянному значению u.

При фиксированном значении u задача состоит в проверке простой гипотезы H_0: $\mu = \mu_0$ против простой альтернативы H_1': $\mu = \mu_1$, $\sigma = \sigma_1$. НКО согласно (22.6) дается в этом случае неравенством

$$\frac{L(x \mid H_0)}{L(x \mid H_1)} \leqslant k_\alpha,$$

которое сводится после упрощений к условию

$$\bar{x}(\mu_1 - \mu_0) \geqslant C(\mu_0, \mu_1, \sigma^2, \sigma_1^2, k_\alpha, u), \tag{23.37}$$

где C — константа, зависящая от x только через u. Отсюда следует, что НКО состоит из больших значений $\bar{x}$, если $\mu_1 - \mu_0 > 0$,

и из малых значений $\bar{x}$, если $\mu_1 - \mu_0 < 0$, причем это верно независимо от значений σ^2, σ_1^2 и $|\mu_1 - \mu_0|$. Таким образом, мы имеем общую НКО для класса альтернатив H_1: $\mu = \mu_1$ в каждой из односторонних ситуаций $\mu_1 > \mu_0$ или $\mu_1 < \mu_0$.

Мы считали значение u фиксированным. Теперь

$$u = \sum (x - \mu_0)^2 = \sum (x - \bar{x})^2 + n(\bar{x} - \mu_0)^2 = \tag{22.38}$$

$$= \sum (x - \bar{x})^2 \left\{ 1 + \frac{n(\bar{x} - \mu_0)^2}{\sum (x - \bar{x})^2} \right\}. \tag{23.39}$$

Поскольку НКО для фиксированного значения u состоит из крайних значений $\bar{x}$, из (23.38) следует, что НКО состоит из малых значений $\sum (x - \bar{x})^2$, что согласно (23.39) соответствует большим значениям величины

$$\frac{t^2}{n-1} = \frac{n(\bar{x} - \mu_0)^2}{\sum (x - \bar{x})^2}, \tag{23.40}$$

где t^2 — квадрат t-статистики Стьюдента, распределение которой было выведено в примере 11.8. Из упражнения 23.7 следует, что статистика t, имея распределение, не зависящее от параметра σ^2, не зависит от полной достаточной статистики u для σ^2. Вспоминая теперь о знаке $\bar{x}$, мы получаем окончательно, что безусловный РНМ подобный критерий для H_0 против H_1 состоит в том, что надо отвергнуть наибольшие или наименьшие 100α процентов распределения t в зависимости от того, что больше: μ_1 или μ_0.

Как мы уже видели, распределение t не зависит от σ^2. Однако мощность РНМ подобного критерия зависит от σ^2, так как u не является достаточной статистикой для σ^2, когда H_0 не выполняется. Поскольку любая подобная область для H_0 состоит из долей α каждого контура, соответствующего постоянному значению u, и распределение на любом таком контуре зависит от σ^2 при гипотезе H_1, то не существует ни одной подобной области для H_0, мощность которой не зависела бы от σ^2. Этот результат был впервые получен Данцигом (1940).

Пример 23.8

Для двух нормальных распределений со средними μ и $\mu + \theta$ и общей дисперсией σ^2 проверить гипотезу

$$H_0 : \theta = \theta_0 \quad (=0 \text{ без ограничения общности})$$

против

$$H_1 : \theta = \theta_1$$

на основе независимых выборок объемов n_1 и n_2 со средними $\bar{x}_1$ и $\bar{x}_2$. Обозначим

$$\left.\begin{aligned} n = n_1 + n_2, \quad n\bar{x} = n_1\bar{x}_1 + n_2\bar{x}_2, \\ s^2 = \sum_{i=1}^{2}\sum_{j=1}^{n_i}(x_{ij} - \bar{x})^2 = \sum\sum x_{ij}^2 - n\bar{x}^2. \end{aligned}\right\} \tag{23.41}$$

Обе гипотезы являются сложными с двумя степенями свободы. Когда выполнена H_0 (но не в других случаях), пара статистик $(\bar{x}, s^2)$ достаточна для мешающих параметров (μ, σ^2) и, как следует из **23.10**, полна. Таким образом, все подобные области для H_0 удовлетворяют (23.13), и при фиксированых $(\bar{x}, s^2)$ мы должны проверить простую гипотезу H_0: $\theta = 0$ против гипотезы H_1': $\theta = \theta_1$, $\mu = \mu_1$, $\sigma = \sigma_1$.

Исходная гипотеза H_1 представляет собой класс гипотез H_1', соответствующих всем μ_1 и σ_1.

НКО, полученная с помощью (22.6), после упрощений принимает вид

$$\bar{x}_2\theta_1 \leqslant g_\alpha,$$

где g_α — постоянная, зависящая от всех параметров и от $\bar{x}$ и s^2, но не зависящая от наблюдений никаким другим способом. Следовательно, при фиксированных $\bar{x}$ и s^2 НКО определяется крайними значениями $\bar{x}_2$, имеющими знак, противоположный знаку θ_1, причем это справедливо при любых значениях других параметров. Из (23.41) следует, что для любых фиксированных $(\bar{x}, s^2)$ НКО будет состоять из больших значений величины $\frac{(\bar{x}_1 - \bar{x}_2)^2}{s^2}$ или, что эквивалентно, монотонно возрастающей функции от нее

$$\frac{(\bar{x}_1 - \bar{x}_2)^2}{\sum(x_1 - \bar{x}_1)^2 + \sum(x_2 - \bar{x}_2)^2} = \frac{t^2}{(n-2)} \cdot \frac{n}{n_1 n_2}. \tag{23.42}$$

(23.42) есть определение обычной t-статистики Стьюдента в данной задаче, с которой мы уже имели дело в **21.12** как с задачей интервального оценивания. В силу упражнения 23.7 статистика t^2, распределение которой не зависит от параметров μ и σ^2, распределена независимо от полной достаточной статистики $(\bar{x}, s^2)$ для (μ, σ^2). Таким образом, безусловный РНМ подобный критерий для H_0 против H_1 заключается в том, чтобы отвергнуть 100α процентов наибольших или наименьших значений в распределении t в соответствии с тем, положительно θ_1 (или, в общем случае, $\theta_1 - \theta_0$) или отрицательно.

Здесь, как и в предыдущем примере, мощность НКО зависит от (μ, σ^2), поскольку статистики $(\bar{x}, s^2)$ не образуют достаточной системы, если H_0 не выполнена.

Пример 23.9

Проверить сложную гипотезу H_0: $\sigma = \sigma_0$ против H_1: $\sigma = \sigma_1$ в случае распределения

$$dF = \exp\left\{-\left(\frac{x-\theta}{\sigma}\right)\right\}\frac{dx}{\sigma}, \qquad \theta \leqslant x \leqslant \infty, \qquad \sigma > 0.$$

В примере 17.19 было показано, что наименьшее из n наблюдений $x_{(1)}$ достаточно для мешающего параметра θ как при гипотезе H_0, так и при H_1. В силу **23.12** эта статистика полна. Следовательно, любая подобная область содержит долю α каждого контура, соответствующего постоянному значению $x_{(1)}$. При фиксированном $x_{(1)}$ гипотезы H_0 и H_1 становятся простыми. НКО, полученная из (22.6), состоит из точек, удовлетворяющих неравенству

$$\sum_{i=1}^{n} x_i\left(\frac{1}{\sigma_1} - \frac{1}{\sigma_0}\right) \leqslant g_\alpha,$$

где g_α — константа, зависящая от σ_0 и σ_1. Таким образом, при каждом фиксированном $x_{(1)}$ мы получаем НКО вида

$$\left.\begin{aligned} &\sum_{i=1}^{n} x_i \leqslant a_\alpha, \quad \text{если} \quad \sigma_1 < \sigma_0, \\ &\sum_{i=1}^{n} x_i \geqslant b_\alpha, \quad \text{если} \quad \sigma_1 > \sigma_0. \end{aligned}\right\} \tag{23.43}$$

Распределение статистики $\sum x_i$ в (23.43) зависит от $x_{(1)}$. Чтобы преобразовать (23.43) к более удобной для использования форме, заметим, что распределение статистики

$$z = \sum_{i=1}^{n} (x_{(i)} - x_{(1)})$$

не зависит от $x_{(1)}$. (Это следует из полноты и достаточности $x_{(1)}$ — см. упражнение 23.7 ниже.) Поэтому, если мы перепишем (23.43) для фиксированного $x_{(1)}$ в виде

$$\left.\begin{aligned} &z \leqslant c_\alpha, \quad \sigma_1 < \sigma_0, \\ &z \geqslant d_\alpha, \quad \sigma_1 > \sigma_0, \end{aligned}\right\} \tag{23.44}$$

где $c_\alpha = a_\alpha - nx_{(1)}$, $d_\alpha = b_\alpha - nx_{(1)}$, то мы получим в левой части (23.43) статистику, крайние значения которой для каждого фиксированного $x_{(1)}$ определяют НКО и распределение которой не зависит от $x_{(1)}$. Таким образом, (23.44) дает безусловную НКО в каждой из односторонних ситуаций $\sigma_1 < \sigma_0$ и $\sigma_1 > \sigma_0$, т. е. мы получаем обычную пару РНМ критериев.

Заметим, что в этом примере достаточность $x_{(1)}$ как при H_0, так и при H_1 делает мощность РНМ критерия не зависящей от параметра расположения θ.

23.21 В примерах 23.7 и 23.8 было дано тщательное обоснование двух стандартных процедур, относящихся к средним нормального распределения. В упражнениях 23.13 и 23.14 с помощью таких же рассуждений обосновываются две стандартные процедуры, относящиеся к дисперсиям, причем в обоих случаях находится пара РНМ подобных односторонних критериев. К сожалению, не все задачи проверки гипотез, связанных с нормальным распределением, решаются столь же успешно. Так, наиболее трудная из этих задач, задача о двух средних, которая детально рассматривалась в главе 21, как показывает следующий пример, не может быть решена с помощью использованных выше методов.

Пример 23.10

Для двух нормальных распределений со средними и дисперсиями (θ, σ_1^2) и $(\theta + \mu, \sigma_2^2)$ проверить гипотезу H_0: $\mu = 0$ на основании двух независимых выборок, состоящих из n_1 и n_2 наблюдений.

Если выполнена гипотеза H_0, то выборочные средние и дисперсии $(\bar{x}_1, \bar{x}_2, s_1^2, s_2^2) = \boldsymbol{t}$ образуют систему из четырех совместно достаточных статистик для трех параметров θ, σ_1^2, σ_2^2, значения которых не определяются гипотезой H_0. Используя (23.21), можно также показать, что они дают минимальную достаточную систему (см. Леман и Шеффе (1950)). Однако статистика $\boldsymbol{t}$ не будет ограниченно полной, поскольку $\bar{x}_1$, $\bar{x}_2$ распределены нормально независимо от s_1^2, s_2^2 и друг от друга, так что любая ограниченная нечетная функция от $(\bar{x}_1 - \bar{x}_2)$ будет иметь нулевое математическое ожидание. Мы, следовательно, не можем использовать (23.13) для нахождения всех подобных областей, хотя согласно **23.8** области, удовлетворяющие (23.13), конечно, будут подобными. Далее, из того факта, что функция правдоподобия содержит четыре компоненты $\boldsymbol{t}$ и не содержит других функций от наблюдений, немедленно следует, что любая область, целиком состоящая из долей α каждой поверхности, соответствующей постоянному значению $\boldsymbol{t}$, будет иметь одну и ту же вероятностную меру в выборочном пространстве *при любых значениях* μ и будет, следовательно, неэффективной критической областью с мощностью, в точности равной ее размеру. Этот неприятный аспект хорошо знакомого и полезного свойства нормального распределения был отмечен Ватсоном (1957a).

Для приведенной задачи не существует никакой полезной точной нерандомизованной подобной области (см. Линник (1964)). Если мы согласны использовать асимптотически подобные области, можно применить метод Уэлча, изложенный в **21.25**, в качестве одного из способов интервального оценивания; если же допустить введение рандомизации, то можно прибегнуть к методу Шеффе (**21.15—22**). Связь терминологии доверительных интервалов с терминологией теории критериев рассматривается в **23.26** ниже.

23.22 Из рассмотрений пункта **23.20** и примеров 23.8—23.10 становится ясно, что если имеется полная достаточная статистика для мешающего параметра, то задача выбора наиболее мощного критерия для сложной гипотезы значительно упрощается, если мы ограничим наш выбор подобными областями. Однако на этом можно и проиграть: для определенных альтернатив может существовать неподобный критерий, удовлетворяющий (23.6) и обладающий большей мощностью, чем наиболее мощный подобный критерий.

Леман и Стейн (1948) рассмотрели эту задачу для сложных гипотез из примера 23.7 и упражнения 23.13. В первом случае, где проверяется гипотеза о среднем нормального распределения, они нашли, что при $\alpha \geqslant 1/2$ не существует неподобного критерия более мощного, чем t-критерий Стьюдента, каковы бы ни были μ_1 и σ_1, однако при $\alpha < 1/2$ (как всегда бывает на практике) существует более мощная критическая область, имеющая вид

$$\sum_i \{x_i - c_\alpha(\mu_1, \sigma_1)\}^2 \leqslant k_\alpha(\mu_1, \sigma_1). \tag{23.45}$$

Аналогично, в случае проверки гипотезы о дисперсии нормального распределения (упражнение 23.13 ниже) они нашли, что при $\sigma_1 > \sigma_0$ не существует более мощного неподобного критерия, однако при $\sigma_1 < \sigma_0$ область

$$\sum_i (x_i - \mu_1)^2 \leqslant k_\alpha \tag{23.46}$$

является более мощной, чем наилучшая подобная критическая область.

Таким образом, если мы достаточно ограничим класс альтернатив H_1, то мы сможем иногда, снимая требование подобия, улучшить мощность критерия, делая в то же время ошибку первого рода в среднем меньшей α. На практике, однако, это не служит очень сильным аргументом против использования подобных областей, и именно потому, что мы редко находимся в ситуации, когда диапазон альтернатив к сложной гипотезе ограничен.

Смещение критериев

23.23 В предыдущей главе (**22.26—28**) мы кратко рассмотрели задачу проверки простой гипотезы H_0 против двустороннего класса альтернатив, когда РНМ критерия, вообще говоря, не существует. Теперь мы рассмотрим эту задачу с другой точки зрения, хотя, как мы увидим, двусторонняя природа альтернативной гипотезы будет несущественной для наших рассуждений.

Пример 23.11

Рассмотрим снова задачу, с которой мы имели дело в примерах 22.2, 22.3 и в **22.27**, относящуюся к проверке гипотезы о среднем μ нормального распределения с известной дисперсией, которую мы для удобства полагаем равной единице. Предположим, что мы ограничиваемся критериями, основанными на распределении выборочного среднего $\bar{x}$; это можно сделать согласно **23.3**, поскольку статистика $\bar{x}$ достаточна. Обобщая (22.55), рассмотрим область размера α, определенную неравенствами

$$\bar{x} \leqslant a_{\alpha_1}, \qquad \bar{x} \geqslant b_{\alpha_2}, \tag{23.47}$$

где $\alpha_1 + \alpha_2 = \alpha$ и α_1 теперь не обязательно равно α_2. Значения a и b определяются, как и в (22.15), формулами

$$a_\alpha = \mu_0 - d_\alpha/n^{1/2}, \qquad b_\alpha = \mu_0 + d_\alpha/n^{1/2}$$

и

$$G(-d_\alpha) = \int_{-\infty}^{-d_\alpha} (2\pi)^{-1/2} \exp\left(-\tfrac{1}{2}y^2\right) dy = \alpha.$$

Без ограничения общности мы будем считать $d_\alpha > 0$.

Так же, как и в (22.56), находим мощность критической области (23.47):

$$P = G\{n^{1/2}\Delta - d_{\alpha_2}\} + G\{-n^{1/2}\Delta - d_{\alpha_1}\}, \tag{23.48}$$

где $\Delta = \mu_1 - \mu_0$.

Будем рассматривать мощность (23.48) как функцию Δ. Ее первые две производные равны

$$P' = \left(\frac{n}{2\pi}\right)^{1/2} \left[\exp\left\{-\tfrac{1}{2}(n^{1/2}\Delta - d_{\alpha_2})^2\right\} - \exp\left\{-\tfrac{1}{2}(n^{1/2}\Delta + d_{\alpha_1})^2\right\}\right] \tag{23.49}$$

и

$$P'' = \frac{n}{(2\pi)^{1/2}} \left[(d_{\alpha_2} - n^{1/2}\Delta)\exp\left\{-\tfrac{1}{2}(n^{1/2}\Delta - d_{\alpha_2})^2\right\} + (n^{1/2}\Delta + d_{\alpha_1})\exp\left\{-\tfrac{1}{2}(n^{1/2}\Delta + d_{\alpha_1})^2\right\}\right], \tag{23.50}$$

Согласно (23.49) $P' = 0$, только если

$$\Delta = (d_{\alpha_2} - d_{\alpha_1})/(2n^{1/2}). \tag{23.51}$$

В случае выполнения (23.51) из (23.50) получаем

$$P'' = \frac{n}{(2\pi)^{1/2}}(d_{\alpha_1} + d_{\alpha_2}) \exp\left\{-\frac{1}{2}(n^{1/2}\Delta + d_{\alpha_1})^2\right\}. \tag{23.52}$$

Поскольку мы считаем d_α положительным, то $P'' > 0$ в стационарной точке, которая, следовательно, дает минимум. Согласно (23.51) он достигается при $\Delta = 0$, только когда $\alpha_1 = \alpha_2$. Этот случай был рассмотрен в **22.27**. В других случаях единственный минимум находится в некоторой точке μ_m, где

$$\mu_m > \mu_0, \quad \text{если} \quad \alpha_1 > \alpha_2, \quad \mu_m < \mu_0, \quad \text{если} \quad \alpha_1 < \alpha_2$$

23.24 Из примера 23.11 вытекает, что (исключая случай $\alpha_1 = \alpha_2$) в классе альтернатив H_1 существуют значения μ, для которых вероятность отвергнуть H_0 меньше, чем для μ из H_0. (Заметим, что если бы мы рассматривали односторонний класс альтернатив, скажем $\mu_1 > \mu_0$, то такая ситуация возникла бы в случае неправильного расположения критической области, скажем $\bar{x} \leqslant a_\alpha$.) Ясно, что критерий, отвергающий гипотезу с большей вероятностью, когда она верна, чем когда она неверна, нежелателен. В самом деле, мы могли бы улучшить такой критерий, отвергая гипотезу с вероятностью α при помощи таблицы случайных чисел. Мощность такой процедуры всегда была бы равна α.

Теперь можно сделать следующее обобщение. Будем говорить, что критическая область w размера α для гипотезы H_0: $\theta = \theta_0$ против простой альтернативы H_1: $\theta = \theta_1$ дает *несмещенный* критерий для H_0 против H_1, если ее мощность удовлетворяет условию

$$\mathbf{P}\{\boldsymbol{x} \in w \mid \theta_1\} \geqslant \alpha. \tag{23.53}$$

В противном случае область w и соответствующий ей критерий называются *смещенными**). Если H_1 — сложная гипотеза, а (23.53) выполняется для каждого элемента H_1, то w называется несмещенной критической областью против H_1. Следует заметить, что для несмещенности не требуется, чтобы функция мощности имела регулярный минимум в точке θ_0, как это имело место в примере 23.11 для $\alpha_1 = \alpha_2$, хотя на практике часто бывает именно так. Рис. 22.3 на странице 244 показывает вид

*) Данное использование термина «смещенный» никак не связано с его употреблением в теории оценивания. К счастью, в контексте редко возникает возможность смешения двух указанных значений этого слова.

функций мощности несмещенного критерия (сплошная линия) и двух смещенных критериев.

Если не существует несмещенного критерия, то можно попытаться найти «локально несмещенный типа М» критерий (Кришнан, 1966), средняя мощность которого $\geqslant \alpha$ в окрестности H_0.

Свойство несмещенности настолько привлекательно с интуитивной точки зрения, что естественно ограничиться только несмещенными критериями и искать РНМ несмещенные (РНМН) критерии, которые могут существовать даже при двусторонних альтернативах, когда, вообще говоря, не существует РНМ критериев, если класс рассматриваемых критериев не ограничен. Так, в примере 23.11 «симметричные» критерии, основанные на $\bar{x}$, являются, как легко видеть, РНМН в классе рассматриваемых в этом примере критериев. В действительности, как будет показано в **23.33**, они являются РНМН среди всех критериев для проверки гипотезы H_0.

Пример 23.12

Мы перенесли в упражнение 23.13 доказательство того, что для нормального распределения со средним μ и дисперсией σ^2 статистика $z = \sum_{i=1}^{n} (x_i - \bar{x})^2$ дает пару односторонних РНМ подобных критериев для проверки гипотезы H_0: $\sigma = \sigma_0^2$, причем соответствующие НКО имеют вид

$$z \geqslant a_\alpha, \text{ если } \sigma_1 > \sigma_0, \quad z \leqslant b_\alpha, \text{ если } \sigma_1 < \sigma_0.$$

Теперь мы рассмотрим двустороннюю альтернативную гипотезу

$$H_1\colon \sigma^2 \neq \sigma_0^2.$$

Согласно **22.18** в этом случае не существует РНМ критерия для H_0 против H_1, однако интуитивно кажется разумным использовать статистику z, расщепляя критическую область на две равные части, соответствующие «хвостам» ее распределения, в надежде получить несмещенность, как в примере 23.11. Таким образом, мы отвергаем H_0, если

$$z \geqslant a_{\alpha/2} \text{ или } z \leqslant b_{\alpha/2}.$$

Эта критическая область, конечно, подобна, так как распределение z не зависит от мешающего параметра μ. Поскольку z/σ^2 имеет распределение χ^2 с $n-1$ степенями свободы как при H_0, так и при H_1, то мы получаем

$$a_{\alpha/2} = \sigma_0^2 \chi^2_{1-\alpha/2}, \quad b_{\alpha/2} = \sigma_0^2 \chi^2_{\alpha/2},$$

где χ^2_α есть 100α-процентная точка распределения χ^2. Если выполнена гипотеза H_1, то указанному распределению подчиняется величина z/σ_1^2, так что H_0 отвергается, когда

$$\frac{z}{\sigma_1^2} \geqslant \frac{\sigma_0^2}{\sigma_1^2}\chi_{1-\alpha/2} \quad \text{или} \quad \frac{z}{\sigma_1^2} \leqslant \frac{\sigma_0^2}{\sigma_1^2}\chi^2_{\alpha/2}.$$

Мощность критерия против любого альтернативного значения σ_1^2 равна сумме вероятностей этих двух событий, т. е. вероятности того, что случайная величина, распределенная по закону χ^2, примет значение вне интервала, определяемого $100\left(\frac{1}{2}\alpha\right)$- и $100\left(1-\frac{1}{2}\alpha\right)$-процентными точками, умноженными каждая на σ_0^2/σ_1^2. На рис. 23.1 приведена функция мощности, полученная с помощью описанных вычислений Нейманом и Пирсоном (1936b) для $n=3$ и $\alpha=0{,}02$. Мы видим, что когда $0{,}5 < \sigma_1^2/\sigma_0^2 < 1$, мощность меньше α, так что критерий смещен.

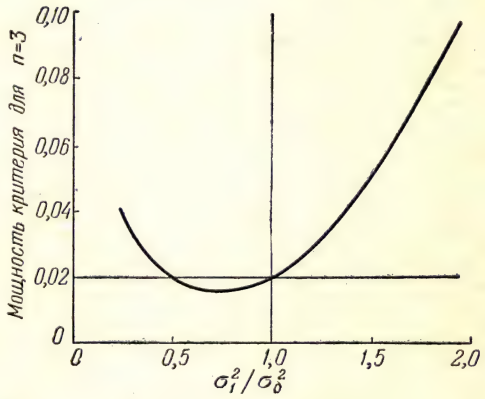


Рис. 23.1. Функция мощности критерия для дисперсии нормального распределения (см. текст).

Посмотрим теперь, можно ли путем перераспределения критической области между «хвостами» распределения z устранить смещение. Пусть критическая область имеет вид

$$z \geqslant a_{1-\alpha_1} \quad \text{или} \quad z \leqslant b_{\alpha_2}$$

где $\alpha_1 + \alpha_2 = \alpha$. Как и раньше, мощность критерия равна вероятности того, что распределенная по закону χ^2 с $n-1$ степенями свободы случайная величина y_{n-1} примет значение вне интервала, определяемого $100\alpha_2$- и $100(1-\alpha_1)$-процентными точками, умноженными каждая на константу $\theta = \sigma_0^2/\sigma_1^2$. Обозначая F функцию распределения y_{n-1}, находим

$$P = F(\theta\chi^2_{\alpha_2}) + 1 - F(\theta\chi^2_{1-\alpha_1}). \qquad (23.54)$$

Рассматриваемое как функция от θ, выражение (23.54) дает функцию мощности. Выберем теперь α_1 и α_2 так, чтобы эта функция мощности имела регулярный минимум в точке $\theta = 1$,

в которой она равна размеру критерия. Дифференцируя (23.54), получаем

$$P' = \chi^2_{\alpha_2} f(\theta\chi^2_{\alpha_2}) - \chi^2_{1-\alpha_1} f(\theta\chi^2_{1-\alpha_1}), \tag{23.55}$$

где f — плотность y_{n-1}. Требуя, чтобы производная была равна нулю при $\theta = 1$, получаем

$$\chi^2_{\alpha_2} f(\chi^2_{\alpha_2}) = \chi^2_{1-\alpha_1} f(\chi^2_{1-\alpha_1}). \tag{23.56}$$

Подставляя в (23.56) функцию плотности

$$f(y) \propto e^{-y/2} y^{(n-3)/2}, \tag{23.57}$$

получаем окончательно следующее условие несмещенности:

$$\left\{\frac{\chi^2_{1-\alpha_1}}{\chi^2_{\alpha_2}}\right\}^{(n-1)/2} = \exp\left\{\frac{1}{2}(\chi^2_{1-\alpha_1} - \chi^2_{\alpha_2})\right\}. \tag{23.58}$$

Значения α_1 и α_2, удовлетворяющие (23.58), дают критерий, функция мощности которого имеет при $\theta = 1$ нулевую производную. Чтобы проверить, является ли полученный критерий строго несмещенным, запишем (23.55), используя (23.57) и (23.58), в виде

$$P' = c\theta^{(n-3)/2}\chi^{n-1}_{\alpha_2} \exp\left(-\frac{1}{2}\chi^2_{\alpha_2}\right)\left[\exp\left\{\frac{1}{2}\chi^2_{\alpha_2}(1-\theta)\right\} - \exp\left\{\frac{1}{2}\chi^2_{1-\alpha_1}(1-\theta)\right\}\right], \tag{23.59}$$

где c — положительная константа. Поскольку $\chi^2_{1-\alpha_1} > \chi^2_{\alpha_2}$, из (23.59) получаем

$$P' \begin{cases} < 0, & \theta < 1, \\ = 0, & \theta = 1, \\ > 0, & \theta > 1, \end{cases} \tag{23.60}$$

откуда следует, что критерий с α_1 и α_2, определенными соотношением (23.58), несмещен в строгом смысле, так как функция мощности монотонно убывает при изменении θ от 0 до 1 и монотонно возрастает при изменении θ от 1 до ∞.

Таблицы значений $\chi^2_{\alpha_2}$ и $\chi^2_{1-\alpha_1}$, удовлетворяющих (23.58), даны Рамачандраном (1958) для $\alpha = 0{,}05$ и $n - 1 = 2(1)\ 8(2)\ 24,\ 30,\ 40,\ 60$; другие таблицы описаны в **20.21** (3) в терминах доверительных интервалов (см. **23.26** относительно связи с критериями). В таблице 23.1 сравниваются некоторые значения, полученные Рамачандраном, с полученными из *Biometrika Tables* значениями для рассмотренного нами критерия с равными «хвостами».

Мы видим, что разности между границами относительно велики при малых n. При возрастании n разность между нижними границами растет, а между верхними убывает. При

Таблица 23.1

Границы для статистики $\sum(x-\bar{x})^2/\sigma_0^2$, за пределами которых отвергается гипотеза $H_0: \sigma^2=\sigma_0^2$ ($\alpha=0{,}05$)

Число степеней свободы ($n-1$)	Границы для несмещенного критерия	Границы для критерия с равными «хвостами»	Разности
2	(0,08, 9,53)	(0,05, 7,38)	(0,03, 2,15)
5	(0,99, 14,37)	(0,83, 12,83)	(0,16, 1,54)
10	(3,52, 21,73)	(3,25, 20,48)	(0,27, 1,25)
20	(9,96, 35,23)	(9,59, 34,17)	(0,37, 1,06)
30	(17,21, 47,96)	(16,79, 46,98)	(0,42, 0,98)
40	(24,86, 60,32)	(24,43, 59,34)	(0,43, 0,98)
60	(40,93, 84,23)	(40,48, 83,30)	(0,45, 0,93)

$n-1=60$ обе разности составляют свыше 1% от значений соответствующих границ.

Вопрос о том, является ли найденный критерий РНМН, мы отложим до примера 23.14.

Несмещенные критерии и подобные критерии

23.25 Имеется тесная связь между несмещенностью и подобием, благодаря которой часто можно прийти к наилучшему несмещенному критерию, исходя из анализа подобных областей.

Пусть рассматривается гипотеза более общего вида, чем (23.2), а именно гипотеза

$$H_0': \theta_r \leqslant \theta_{r0}, \tag{23.61}$$

которая проверяется против альтернативы

$$H_1: \theta_r > \theta_{r0}. \tag{23.62}$$

Если мы найдем критическую область w, удовлетворяющую (23.6) для всех θ_r из H_0' и для всех значений мешающих параметров θ_s, т. е. удовлетворяющую неравенству

$$P(H_0', \theta_s) \leqslant \alpha \tag{23.63}$$

(где P — функция мощности, дающая вероятность отвергнуть H_0), то критерий, основанный на w, будет иметь, как и раньше, размер α. Если он, кроме того, несмещен, то из (23.53) получим

$$P(H_1, \theta_s) \geqslant \alpha. \tag{23.64}$$

Далее, если функция мощности P непрерывна по θ_r, то из (23.63) и (23.64) получаем, учитывая вид гипотез H_0' и H_1, соотношение

$$P(\theta_{r0}, \theta_s) = \alpha, \tag{23.65}$$

означающее, что w является подобной критической областью для «граничной» гипотезы

$$H_0\colon\ \theta_r = \theta_{r0}.$$

Все несмещенные критерии для H_0' суть подобные критерии для H_0. Поэтому, когда мы ограничиваемся рассмотрением только подобных критериев для H_0 и, применяя изученные ранее методы, находим критерий с оптимальными свойствами, например РНМ подобный критерий, нам достаточно убедиться, что найденный критерий несмещен, чтобы сделать вывод о том, что этот критерий оптимален в классе всех несмещенных критериев для H_0', т. е. является РНМН критерием.

Приведенные рассуждения сохраняют силу также в том случае, когда H_0' утверждает, что параметрическая точка θ_r лежит внутри некоторой области R (которая может состоять из нескольких подобластей) в пространстве параметров, а гипотеза H_1 утверждает, что θ_r лежит в дополнении к R. Если функция мощности непрерывна по θ_r, то критическая область w, несмещенная для H_0', будет подобной для гипотезы H_0, утверждающей, что θ_r лежит на границе R. Если w дает несмещенный критерий для H_0', то она сохранит в классе несмещенных критериев для H_0' любые оптимальные свойства, которые она может иметь, будучи подобным критерием для H_0. Однако в случае двусторонних альтернатив РНМ подобный критерий для H_0 не всегда существует. В этом случае может существовать РНМН критерий, но для его нахождения должны быть использованы другие методы.

Пример 23.13.

Вернемся к гипотезе примера 23.12. Односторонние критические области $z \geqslant a_\alpha$, $z \leqslant b_\alpha$, основанные на статистике z, дают РНМ подобные критерии против односторонних альтернатив. Легко видеть, что каждый из них будет несмещенным при проверке одной из гипотез

$$H_0'\colon\ \sigma^2 \leqslant \sigma_0^2, \qquad H_0''\colon\ \sigma^2 \geqslant \sigma_0^2$$

соответственно против

$$H_1'\colon\ \sigma^2 > \sigma_0^2, \qquad H_1''\colon\ \sigma^2 < \sigma_0^2.$$

Следовательно, они, согласно **23.25**, будут РНМН критериями для рассматриваемых односторонних ситуаций.

Однако в случае двусторонней альтернативы $H_1\colon\ \sigma^2 \neq \sigma_0^2$ этим методом уже нельзя показать, что несмещенный критерий, основанный на (23.58), будет РНМН, так как не было показано, что он является РНМ подобным.

Критерии и доверительные интервалы

23.26 Начальная разработка теории несмещенных критериев была в основном проведена в упомянутой в **22.1** серии статей Неймана и Пирсона, а также в статьях Неймана (1935, 1938b), Шеффе (1942a) и Лемана (1947). Многие второстепенные детали этих работ теперь, конечно, устарели, однако их терминология используется еще довольно часто, и поэтому желательно составить «словарь», связывающий ее с современной терминологией в случаях, когда есть расхождения. При этом уместно будет, кроме того, перевести идеи теории проверки гипотез на язык теории доверительных интервалов, как было обещано в **20.20**.

Получив выборку, мы можем задать вопрос: при каких значениях θ выборочная точка $\boldsymbol{x}$ принадлежит области принятия A, дополнительной к критической области размера α, соответствующей некоторому критерию для параметра θ? Объединив эти «приемлемые» значения θ, мы получим доверительный интервал C для θ уровня $1-\alpha$, соответствующий данному критерию, так как θ будет принадлежать C тогда и только тогда, когда $\boldsymbol{x}$ принадлежит A, т. е. с вероятностью $1-\alpha$. Мы использовали этот метод построения доверительных интервалов в **20.3** и фактически во всей главе 20. Таким образом, нет необходимости выводить оптимальные свойства отдельно для критериев и интервалов, поскольку между этими задачами существует взаимно однозначное соответствие.

Свойство критерия		Свойство соответствующего доверительного интервала
Современная терминология	Старая терминология	
РНМ		«Кратчайший» (= наиболее селективный) Несмещенный
Несмещенный РНМН «локально» (т. е. вблизи H_0)	Типа A *) (простая гипотеза H_0, один параметр)	«Короткий» несмещенный
	Типа B *) (сложная гипотеза H_0)	
	Типа C *) (простая гипотеза H_0, два или более параметров)	
РНМН	Типа A_1 *) (простая гипотеза H_0, один параметр)	«Кратчайший» несмещенный
	Типа B_1 *) (сложная гипотеза H_0)	
Несмещенный подобный	Bisimilar **)	

*) При условиях регулярности.

**) Здесь, очевидно, представляет интерес именно английский термин. (*Прим. ред*)

Так, в **20.31** при нахождении с помощью t-распределения Стьюдента доверительного интервала для среднего μ нормального распределения с неизвестной дисперсией было отмечено, что длина этого интервала является случайной величиной, кратной выборочному стандартному отклонению. С другой стороны, в примере 23.7 было показано, что мощность подобного критерия, основанного на t-распределении Стьюдента, зависит от неизвестной дисперсии. Заметим теперь, что мощность критерия есть не что иное, как вероятность отвергнуть ложную гипотезу, т. е. в терминах доверительных интервалов — вероятность не покрыть значение μ, отличное от истинного среднего μ_0. Если эта вероятность зависит от неизвестной дисперсии для всех значений μ, то мы, очевидно, не можем заранее задать одновременно и длину доверительного интервала и доверительный уровень. Замеченный нами ранее факт есть, таким образом, следствие другого, полученного позднее.

РНМН критерии для экспоненциального семейства

23.27 Теперь мы приступим к изложению некоторых чрезвычайно содержательных результатов, принадлежащих Леману и Шеффе (1955), которые доказали существование и дали метод построения РНМН критериев для различных параметрических гипотез, относящихся к распределениям экспоненциального семейства (17.86). Мы будем записывать совместное распределение n независимых наблюдений из такого распределения в виде

$$f(\boldsymbol{x}) = D(\boldsymbol{\tau})\, h(\boldsymbol{x}) \exp\left\{ \sum_{j=1}^{r+1} b_j(\boldsymbol{\tau})\, u_j(\boldsymbol{x}) \right\}, \tag{23.66}$$

где $\boldsymbol{x}$ — вектор-столбец $(x_1, \ldots, x_n)$ и $\boldsymbol{\tau}$ — вектор $r+1$ параметров $(\tau_1, \ldots, \tau_{r+1})$. В матричных обозначениях выражение под знаком экспоненты в (23.66) может быть записано в виде $\boldsymbol{u'b}$, где $\boldsymbol{u}$ и $\boldsymbol{b}$ — векторы-столбцы.

Предположим теперь, что нас интересует некоторая линейная форма от параметров

$$\theta = \sum_{j=1}^{r+1} a_{j1} b_j(\boldsymbol{\tau}), \tag{23.67}$$

где $\sum_{j=1}^{r+1} a_{j1}^2 = 1$. Обозначим $\boldsymbol{A}$ ортогональную матрицу (a_{uv}), первый столбец которой состоит из коэффициентов формы (23.67), и перейдем к новому $(r+1)$-мерному вектору параметров (θ, ψ), где ψ — вектор-столбец $(\psi_1, \ldots, \psi_r)$, с помощью уравнения

$$\begin{pmatrix} \theta \\ \psi \end{pmatrix} = \boldsymbol{A'b}. \tag{23.68}$$

Первая строка (23.68) совпадает с (23.67). Предположим теперь, что существует вектор-столбец статистик $\boldsymbol{T} = (s, t_1, \ldots, t_r)$, определяемый соотношением

$$\boldsymbol{T}'\begin{pmatrix}\theta\\ \psi\end{pmatrix} = \boldsymbol{u}'\boldsymbol{b}, \tag{23.69}$$

т. е. предположим, что экспонента в (23.66) может быть записана в виде

$$\theta s(\boldsymbol{x}) + \sum_{j=1}^{r} \psi_j t_j(\boldsymbol{x}).$$

Учитывая (23.68), можно переписать (23.69) в виде

$$\boldsymbol{T}'\begin{pmatrix}\theta\\ \psi\end{pmatrix} = \boldsymbol{u}'\boldsymbol{A}\begin{pmatrix}\theta\\ \psi\end{pmatrix}. \tag{23.70}$$

Это равенство выполняется тождественно по (θ, ψ), поэтому $\boldsymbol{T}' = \boldsymbol{u}'A$ или

$$\boldsymbol{T} = \boldsymbol{A}'\boldsymbol{u}. \tag{23.71}$$

Сравнивая (23.71) с (23.68), мы видим, что каждая компонента $\boldsymbol{T}$ так же выражается через $u_j(\boldsymbol{x})$, как соответствующая компонента (θ, ψ) через $b_j(\tau)$. В частности, первая компонента согласно (23.67) имеет вид

$$s(x) = \sum_{j=1}^{r+1} a_{j1} u_j(\boldsymbol{x}), \tag{23.72}$$

тогда как $t_j(\boldsymbol{x})$, $j = 1, 2, \ldots, r$, ортогональны к $s(\boldsymbol{x})$.

Отметим, что условие $\sum_{j=1}^{r+1} a_{j1}^2 = 1$ не стесняет нас при проверке гипотезы о параметре θ, определенном формулой (23.67), так как для его выполнения требуется лишь изменить постоянный множитель и соответственно подправить гипотезу.

23.28 Таким образом, если мы можем с помощью приема предыдущего пункта свести задачу проверки гипотезы (обычно через посредство достаточных статистик) к стандартной задаче о параметре θ в распределении

$$f(\boldsymbol{x} \mid \theta, \psi) = C(\theta, \psi)\, h(\boldsymbol{x}) \exp\left\{\theta s(\boldsymbol{x}) + \sum_{i=1}^{r} \psi_i t_i(\boldsymbol{x})\right\}, \tag{23.73}$$

то мы можем воспользоваться результатами, резюмированными в **23.10**: при фиксированном гипотетическом значении θ r-компонентный вектор $\boldsymbol{t} = (t_1, \ldots, t_r)$ будет полной достаточной статистикой для r-компонентного параметра $\psi = (\psi_1, \ldots, \psi_r)$, и мы приходим к задаче использования s и $\boldsymbol{t}$ для проверки различных сложных гипотез, касающихся θ, при мешающем параметре ψ.

Простые гипотезы являются частным случаем этой ситуации, соответствующим $r = 0$, т. е. отсутствию мешающих параметров.

23.29 Для решения поставленной задачи нам понадобится более общая формулировка леммы Неймана — Пирсона из **22.10**. Пусть $f(\boldsymbol{x}|\theta)$ — плотность, а θ_i — некоторое подмножество допустимых значений вектора параметров θ $(i = 1, 2, \ldots, k)$. Пусть θ_i^0 обозначает отдельный элемент θ_i, а θ^* — отдельное значение θ. Вектор $\boldsymbol{u}_i(\boldsymbol{x})$ достаточен для θ, когда θ принадлежит θ_i, и имеет распределение $g_i(\boldsymbol{u}_i|\theta_i)$. Вследствие факторизации правдоподобия в случае достаточности условная плотность $f(\boldsymbol{x}|\,\theta_i)$ при фиксированном $\boldsymbol{u}_i$ не зависит от θ_i, и мы обозначим ее $f(\boldsymbol{x}|\boldsymbol{u}_i)$. Пусть, кроме того, $l_i(\boldsymbol{x})$ и $m_i(\boldsymbol{u}_i)$ — неотрицательные функции от наблюдений и от $\boldsymbol{u}_i$ соответственно.

Пусть теперь имеется критическая область w, для которой

$$\int_w \{l_i(\boldsymbol{x}) f(\boldsymbol{x}\,|\,\boldsymbol{u}_i)\}\, d\boldsymbol{x} = \alpha_i. \tag{23.74}$$

Поскольку произведение в фигурных скобках неотрицательно, оно может рассматриваться как плотность некоторого распределения, и мы можем сказать, что условный размер w при фиксированном $\boldsymbol{u}_i$ равен α_i относительно этого распределения. Положим теперь

$$\begin{aligned}\beta_i = \alpha_i \int m_i(\boldsymbol{u}_i)\, g_i(\boldsymbol{u}_i\,|\,\theta_i^0)\, d\boldsymbol{u}_i = \\ = \int_w l_i(\boldsymbol{x})\, m_i(\boldsymbol{u}_i) \Big\{ \int f(\boldsymbol{x}\,|\,\boldsymbol{u}_i)\, g_i(\boldsymbol{u}_i\,|\,\theta_i^0)\, d\boldsymbol{u}_i \Big\}\, d\boldsymbol{x} = \\ = \int_w \{l_i(\boldsymbol{x})\, m_i(\boldsymbol{u}_i) f(\boldsymbol{x}\,|\,\theta_i^0)\}\, d\boldsymbol{x}. \end{aligned} \tag{23.75}$$

Произведение в фигурных скобках снова будет плотностью, которую мы обозначим $p(\boldsymbol{x}\,|\,\theta_i^0)$. Чтобы проверить простую гипотезу о том, что имеет место $p(\boldsymbol{x}\,|\,\theta_i^0)$, против простой альтернативы $f(\boldsymbol{x}|\theta^*)$, мы можем использовать (22.6) и получить НКО w размера β_i, состоящую из точек, удовлетворяющих соотношению

$$f(\boldsymbol{x}\,|\,\theta^*)/p(\boldsymbol{x}\,|\,\theta_i^0) \geqslant c_i(\beta_i), \tag{23.76}$$

где c_i — неотрицательная константа. (23.76) выполняется при каждом значении i, поэтому для проверки сложной гипотезы о том, что имеет место какое-то из распределений $p(\boldsymbol{x}\,|\,\theta_i^0)$ $(i = 1, 2, \ldots, k)$, нужно, чтобы для w были выполнены все k неравенств (23.76). Если мы теперь напишем $km_i(\boldsymbol{u}_i)/c_i(\beta_i)$ вместо $m_i(\boldsymbol{u}_i)$ в определении $p(\boldsymbol{x}\,|\,\theta_i^0)$, что можно сделать, поскольку функция $m_i(\boldsymbol{u}_i)$ произвольна, то мы получим, суммируя

неравенства (23.76) для $i = 1, 2, \ldots, k$, необходимое и достаточное условие для НКО:

$$f(\boldsymbol{x} \mid \theta^*) \geqslant \sum_{i=1}^{k} l_i(\boldsymbol{x})\, m_i(\boldsymbol{u}_i) f(\boldsymbol{x} \mid \theta_i^0). \tag{23.77}$$

Это и есть требуемое обобщение. (22.6) получается отсюда как частный случай при $k = 1$, $l_1(\boldsymbol{x}) = k_\alpha$ (константа), $m_1(u_1) \equiv 1$. При рассмотрении сложных гипотез (23.77) будет играть ту же роль, что и (22.6) в случае простых гипотез.

Односторонние альтернативы

23.30 Возвращаясь к (23.73), мы рассмотрим задачу проверки гипотезы

$$H_0^{(1)}: \ \theta \leqslant \theta_0,$$

против

$$H_1^{(1)}: \ \theta > \theta_0,$$

которая уже обсуждалась в общей постановке в **23.25**. Теперь, когда мы имеем дело с экспоненциальным семейством (23.73), мы сможем показать, что всегда имеется РНМН критерий для $H_0^{(1)}$ против $H_1^{(1)}$. Согласно **23.25**, если имеется несмещенная критическая область размера α для $H_0^{(1)}$ против $H_1^{(1)}$, то она будет подобной областью для проверки гипотезы $\theta = \theta_0$.

Рассмотрим проверку простой гипотезы

$$H_0: \ \theta = \theta_0, \qquad \psi = \psi^0,$$

против простой альтернативы

$$H_1: \ \theta = \theta^* > \theta_0, \qquad \psi = \psi^*.$$

Применим результат пункта **23.29**. Полагая $k = 1$, $l_1(\boldsymbol{x}) \equiv 1$, $\alpha_1 = \alpha$, $\theta = (\theta, \psi)$, $\theta_1 = (\theta_0, \psi)$, $\theta^* = (\theta^*, \psi^*)$, $\theta_1^0 = (\theta_0, \psi^0)$, $\boldsymbol{u}_1 = \boldsymbol{t}$, получаем в силу (23.77) и (22.73), что НКО для проверки H_0 против H_1 имеет вид

$$\frac{C(\theta^*, \psi^*) \exp\left\{\theta^* s(\boldsymbol{x}) + \sum_{i=1}^{r} \psi_i^* t_i(\boldsymbol{x})\right\}}{C(\theta_0, \psi^0) \exp\left\{\theta_0 s(\boldsymbol{x}) + \sum_{i=1}^{r} \psi_i^0 t_i(\boldsymbol{x})\right\}} \geqslant m_1(\boldsymbol{t}) \tag{23.78}$$

или

$$s(\boldsymbol{x})(\theta^* - \theta_0) \geqslant c_\alpha(\boldsymbol{t}, \theta^*, \theta_0, \psi^*, \psi^0). \tag{23.79}$$

Мы видим, что c_α не зависит от ψ. Действительно, так как статистика $\boldsymbol{t}$ достаточна для параметра ψ при гипотезе H_0, то значение c_α при заданном $\boldsymbol{t}$ не зависит от ψ^0, ψ^*. Далее, из (23.79) вытекает, что если знак $(\theta^* - \theta_0)$ не меняется, то НКО состоит из

верхних 100α процентов распределения $s(\boldsymbol{x})$ при заданном θ_0. Таким образом, мы получаем НКО для $\theta = \theta_0$ против $\theta > \theta_0$, дающую РНМ критерий. Мощность этого РНМ критерия не может быть меньше мощности рандомизованного критерия против $\theta > \theta_0$, не учитывающего наблюдений. Последний же имеет мощность, равную своему размеру α. Отсюда следует, что наш РНМ критерий несмещен для $\theta > \theta_0$, т. е. согласно **23.25** является РНМН. Его размер при $\theta < \theta_0$ не превосходит размера при θ_0, так как критическая область (23.79) имеет минимальную мощность против $\theta < \theta_0$, а следовательно, ее мощность (размер) при $\theta < \theta_0$ меньше α. Таким образом, мы показали, что верхние 100α процентов условного распределения $s(\boldsymbol{x})$ при фиксированном $\boldsymbol{t}$ дают РНМН критерий размера α для $H_0^{(1)}$ против $H_1^{(1)}$.

Двусторонние альтернативы

23.31 Рассмотрим теперь задачу проверки гипотезы

$$H_0^{(2)}\colon\ \theta = \theta_0$$

против

$$H_1^{(2)}\colon\ \theta \neq \theta_0.$$

В предшествующих примерах нам не удалось найти РНМН критерия для двусторонних альтернатив этого типа (см. примеры 23.12 и 23.13). Тем не менее для линейной экспоненциальной формы (23.73) РНМН критерий существует.

Из **23.25** следует, что если функция мощности некоторой критической области непрерывна по θ и несмещена, то она подобна для $H_0^{(2)}$. Далее, для любой области w функция мощности имеет вид

$$P(w \mid \theta) = \int_w f(\boldsymbol{x} \mid \theta,\ \psi)\, d\boldsymbol{x}, \tag{23.80}$$

где f определяется в (23.73). Функция (23.80) непрерывна и дифференцируема под знаком интеграла по θ. Таким образом, для того чтобы критерий, основанный на критической области w, был несмещенным, необходимо, чтобы при любом значении ψ выполнялось условие

$$P'(w \mid \theta_0) = 0. \tag{23.81}$$

Дифференцируя (23.80) под знаком интеграла и учитывая (23.73) и (23.81), получаем условие несмещенности

$$0 = \int_w \left[s(\boldsymbol{x}) + \frac{C'(\theta_0,\ \psi)}{C(\theta_0,\ \psi)} \right] f(\boldsymbol{x} \mid \theta_0,\ \psi)\, d\boldsymbol{x}$$

или

$$\mathbf{M}\{s(\boldsymbol{x})\, c(w)\} = -\alpha C'(\theta_0,\ \psi)/\boldsymbol{C}(\theta_0,\ \psi). \tag{23.82}$$

Поскольку в силу (23.73)

$$1/C(\theta, \psi) = \int h(x) \exp\left\{\theta s(x) + \sum_i \psi_i t_i(x)\right\} dx,$$

то мы получаем

$$\frac{C'(\theta, \psi)}{C(\theta, \psi)} = -\mathbf{M}\{s(x)\}. \tag{23.83}$$

Подставляя (23.83) в (23.82), находим

$$\mathbf{M}\{s(x)\, c(w)\} = \alpha \mathbf{M}\{s(x)\}. \tag{23.84}$$

Беря сначала условное относительно t, а затем безусловное математическое ожидание в (23.84), получаем

$$\mathbf{M}_t[\mathbf{M}\{s(x)\, c(w) - \alpha s(x) \mid t\}] = 0. \tag{23.85}$$

В силу полноты t из (23.85) следует

$$\mathbf{M}\{s(x)\, c(w) - \alpha s(x) \mid t\} = 0. \tag{23.86}$$

Далее, любая подобная область для H_0 удовлетворяет условию

$$\mathbf{M}\{c(w) \mid t\} = \alpha, \tag{23.87}$$

объединяя которое с (23.86) мы получаем соотношение

$$\mathbf{M}\{s^{i-1}(x)\, c(w) \mid t\} = \alpha \mathbf{M}\{s^{i-1}(x) \mid t\} = \alpha_i, \quad i = 1, 2, \tag{23.88}$$

где все математические ожидания берутся в предположении, что $\theta = \theta_0$.

Будем проверять теперь простую гипотезу

$$H_0: \ \theta = \theta_0, \quad \psi = \psi^0$$

против простой альтернативы

$$H_1: \ \theta = \theta^* \neq \theta_0, \quad \psi = \psi^*,$$

применяя результат пункта **23.29** при α_i из (23.88), $k = 2$, $\theta = (\theta, \psi)$, $\theta_1 = \theta_2 = (\theta_0, \psi)$, $\theta^* = (\theta^*, \psi^*)$, $\theta_1^0 = \theta_2^0 = (\theta_0, \psi^0)$, $l_i(x) = s^{i-1}(x)$, $u_1 = u_2 = t$. Согласно (23.77) и (23.73) НКО w для проверки H_0 против H_1 имеет вид

$$\frac{C(\theta^*, \psi^*) \exp\left\{\theta^* s(x) + \sum_{i=1}^{r} \psi_i^* t_i(x)\right\}}{C(\theta_0, \psi^0) \exp\left\{\theta_0 s(x) + \sum_{i=1}^{r} \psi_i^0 t_i(x)\right\}} \geqslant m_1(t) + s(x)\, m_2(t), \tag{23.89}$$

или

$$\exp\{s(x)(\theta^* - \theta_0)\} \geqslant c_1(t, \theta^*, \theta_0, \psi^*, \psi^0) + s(x)\, c_2(t, \theta^*, \theta_0, \psi^*, \psi^0),$$

или

$$\exp\{s(x)(\theta^* - \theta_0)\} - s(x)\, c_2 \geqslant c_1 \tag{23.90}$$

Неравенство (23.90) означает, что $s(x)$ лежит вне некоторого интервала, т. е.

$$s(x) \leqslant v(t), \qquad s(x) \geqslant w(t), \tag{23.91}$$

где $v(t)$ и $w(t)$ $(v(t) < w(t))$ могут зависеть от параметров. Мы покажем, однако, что $v(t)$ и $w(t)$ зависят только от θ_0. Как и раньше, достаточность t для ψ устраняет зависимость v и w от ψ при заданном t, а их независимость от θ^* непосредственно следует из соотношения (23.86), в силу которого при гипотезе H_0 имеем

$$\int_w \{s(x) \mid t\} f\, dx = \alpha \int_W \{s(x) \mid t\} f\, dx. \tag{23.92}$$

Правая часть (23.92), где интегрирование ведется по всему выборочному пространству, очевидно, совсем не зависит от θ^*. Поэтому левая часть тоже не зависит от θ^*, так что НКО w, определенная неравенствами (23.91), зависит, как это и должно быть, только от θ_0. Следовательно, эта НКО дает РНМ критерий для $H_0^{(2)}$ против $H_1^{(2)}$. Для доказательства его несмещенности достаточно повторить рассуждения, приведенные в конце пункта **23.30**. Таким образом, мы установили, что НКО (23.91) дает РНМН критерий для $H_0^{(2)}$ против $H_1^{(2)}$. Чтобы построить этот РНМН критерий, надо взять в качестве критической области дополнение интервала, содержащего $100(1-\alpha)$ процентов условного распределения $s(x)$ при данном t, когда верна гипотеза $H_0^{(2)}$; если эта область несмещена, она дает РНМН критерий размера α.

Гипотезы с конечными интервалами

23.32 Можно рассматривать также проверку гипотезы

$$H_0^{(3)}\colon\ \theta_0 \leqslant \theta \leqslant \theta_1$$

против

$$H_1^{(3)}\colon\ \theta < \theta_0 \quad \text{или} \quad \theta > \theta_1$$

или, наоборот,

$$H_0^{(4)}\colon\ \theta \leqslant \theta_0 \quad \text{или} \quad \theta \geqslant \theta_1$$

против

$$H_1^{(4)}\colon\ \theta_0 < \theta < \theta_1.$$

Рассмотрим проверку двух гипотез

$$H_0'\colon\ \theta = \theta_0, \quad \psi = \psi^0, \qquad H_0''\colon\ \theta = \theta_1, \quad \psi = \psi^1$$

против

$$H_1\colon\ \theta = \theta^*, \quad \psi = \psi^*, \quad \text{где} \quad \theta_0 \neq \theta^* \neq \theta_1.$$

Применяя снова результат пункта **23.29**, на этот раз с $k=2$, $\alpha_1=\alpha_2=\alpha$, $\theta=(\theta, \psi)$, $\theta_1=(\theta_0, \psi)$, $\theta_2=(\theta_1, \psi)$, $\theta^*=(\theta^*, \psi^*)$, $\theta_1^0=(\theta_0, \psi^0)$, $\theta_2^0=(\theta_1, \psi^1)$, $l_i(\boldsymbol{x})\equiv 1$, $\boldsymbol{u}_1=\boldsymbol{u}_2=\boldsymbol{t}$, мы получаем, что НКО w для проверки H_0' или H_0'' против H_1 определяется неравенством

$$f(\boldsymbol{x}\,|\,\theta^*, \psi^*) \geqslant m_1(\boldsymbol{t})\, f(\boldsymbol{x}\,|\,\theta_0, \psi^0) + m_2(\boldsymbol{t})\, f(\boldsymbol{x}|\,\theta_1, \psi^1), \quad (23.93)$$

или, после подстановки $f(\boldsymbol{x})$ из (23.73),

$$H(s) = c_1 \exp\{(\theta_0 - \theta^*)\, s(\boldsymbol{x})\} + c_2 \exp\{(\theta_1 - \theta^*)\, s(\boldsymbol{x})\} < 1, \quad (23.94)$$

где c_1, c_2 могут зависеть от всех параметров и от $\boldsymbol{t}$. Если $\theta_0 < \theta^* < \theta_1$, то условие (23.94) эквивалентно требованию, чтобы значение $s(\boldsymbol{x})$ лежало в некотором интервале

$$v(\boldsymbol{t}) \leqslant s(\boldsymbol{x}) \leqslant w(\boldsymbol{t}). \quad (23.95)$$

С другой стороны, если $\theta^* < \theta_0$ или $\theta^* > \theta_1$, то (23.94) требует, чтобы значение $s(\boldsymbol{x})$ лежало *вне* интервала $(v(\boldsymbol{t}), w(\boldsymbol{t}))$. Доказательство того, что границы интервала могут зависеть только от значений параметров θ_0 и θ_1, а также доказательство несмещенности аналогичны приведенным ранее. Таким образом, мы имеем РНМН критерии для $H_0^{(3)}$ и для $H_0^{(4)}$. Как следует из **23.25**, оба критерия подобны при $\theta=\theta_0$ и $\theta=\theta_1$. Чтобы получить РНМН критерий для $H_0^{(3)}$ (или $H_0^{(4)}$), надо найти интервал, содержащий $100(1-\alpha)$ (или 100α для $H_0^{(4)}$) процентов условного распределения $s(\boldsymbol{x})$ при данном $\boldsymbol{t}$ как при $\theta=\theta_0$, так и при $\theta=\theta_1$. Если область, представляющая собой дополнение интервала (или сам интервал), несмещена, то она дает РНМН критерий для проверки $H_0^{(3)}$ (или $H_0^{(4)}$).

23.23 Теперь мы переходим к некоторым приложениям фундаментальных результатов пунктов **23.30—32**, касающихся РНМН критериев для экспоненциального семейства распределений. Отметим сначала, что в примере 23.11 и упражнениях 22.1—22.3, приведенных выше, РНМН критерии для всех четырех типов гипотез были получены непосредственно из распределения одномерной достаточной статистики, без использования условных распределений, поскольку мешающие параметры отсутствовали.

Пример 23.14

В случае n независимых наблюдений из нормального распределения пара статистик $(\bar{x}, s^2)$ совместно достаточна для (μ, σ^2) и имеет совместное распределение (см. пример 17.17)

$$g(\bar{x}, s^2\,|\,\mu, \sigma^2) \propto \frac{s^{n-3}}{\sigma^n} \exp\left\{-\frac{\sum(x-\mu)^2}{2\sigma^2}\right\}, \quad (23.96)$$

которое может быть записано в виде

$$g \propto C(\mu, \sigma^2) \exp\left\{\left(-\frac{1}{2}\sum x^2\right)\left(\frac{1}{\sigma^2}\right)+\left(\sum x\right)\left(\frac{\mu}{\sigma^2}\right)\right\}, \quad (23.97)$$

представляющем частный случай (23.73). Следуя рассуждениям пункта **23.27**, рассмотрим линейную форму от параметров, входящих в (23.97),

$$\theta = A\left(\frac{1}{\sigma^2}\right)+B\left(\frac{\mu}{\sigma^2}\right), \quad (23.98)$$

где A и B — произвольные константы. Мы выберем A и B, чтобы, пользуясь результатами **23.30—32**, получить РНМН критерии для следующих гипотез.

(1) Положим $A = 1$, $B = 0$ и будем проверять гипотезы о параметре $\theta = 1/\sigma^2$, считая мешающим параметром $\psi = \mu/\sigma^2$. В данном случае $s(\boldsymbol{x}) = -\frac{1}{2}\sum x^2$ и $t(\boldsymbol{x}) = \sum x$. Из (23.97) видно, что существует РНМН критерий для гипотез $H_0^{(1)}$, $H_0^{(2)}$, $H_0^{(3)}$ и $H_0^{(4)}$, касающихся $1/\sigma^2$ (а следовательно, и σ^2), основанный на условном распределении $\sum x^2$ при фиксированной $\sum x$, т. е. на условном распределении $\sum (x-\bar{x})^2$ при фиксированной $\sum x$. Поскольку последние две статистики независимы, можно использовать безусловное распределение $\sum (x-\bar{x})^2$ или безусловное распределение $\sum (x-\bar{x})^2/\sigma^2$, совпадающее с распределением χ^2 с $n-1$ степенями свободы. Гипотеза $H_0^{(2)}$ была рассмотрена в примерах 23.12, 23.13, где были получены РНМ подобный критерий для гипотезы $\theta = \theta_0$ против односторонних альтернатив и несмещенный критерий для $H_0^{(2)}$, основанный на $\sum (x-\bar{x})^2$. Теперь мы получили, что последний критерий является РНМН для $H_0^{(2)}$, тогда как односторонний критерий будет РНМН для $H_0^{(1)}$.

Гентер и Уиткомб (1966) дают графики для нахождения критических значений РНМН критериев для проверки $H_0^{(3)}$ и $H_0^{(4)}$ при $\alpha = 0{,}05$ и $0{,}10$.

(2) Чтобы проверять гипотезы, касающиеся μ, запишем (23.98) в виде $\mu = (\theta\sigma^2 - A)/B$. Зафиксировав значение μ_0 для μ, мы все же не сможем найти значения A и B, при которых этому значению соответствовало бы единственное значение $\theta = \theta_0$ (при неизвестном σ^2), если $\theta_0 \neq 0$. Однако, если $\theta_0 = 0$, мы получаем $\mu_0 = -A/B$. Таким образом, имея РНМН критерии для $H_0^{(1)}$: $\theta \leqslant 0$ и $H_0^{(2)}$: $\theta = 0$, можно получить РНМН критерии для $\mu \leqslant \mu_0$ и $\mu = \mu_0$. Используя (23.71), можно показать, что статистика критерия $s(\boldsymbol{x})\,|\,t$ имеет здесь вид $\left(-\frac{1}{2}\sum x^2\right)A+\left(\sum x\right)B$ при условии, что фиксирована неко-

торая ортогональная функция, например $\left(-\frac{1}{2}\sum x^2\right)B-\left(\sum x\right)A$, что сводится к условному распределению $\sum x$ при фиксированной $\sum x^2$. Ясно, что в этом случае мы не можем получить критерии для гипотез $H_0^{(3)}$ или $H_0^{(4)}$, относящихся к μ.

Критерий для проверки гипотезы $\mu = \mu_0$ против односторонних альтернатив был рассмотрен в примере 23.7, где было показано, что t-критерий Стьюдента, к которому он сводится, является РНМ подобным критерием для $\mu = \mu_0$ против односторонних альтернатив. Теперь мы видим, что этот критерий является также РНМН для $H_0^{(1)}$. Мы видим также, что двусторонний t-критерий Стьюдента с равными «хвостами», несмещенный для $H_0^{(2)}$ против $H_1^{(2)}$, будет РНМН критерием для $H_0^{(2)}$.

Пример 23.15

Рассмотрим k независимых выборок объема n_i $(i = 1, 2, \ldots, k)$ из нормальных распределений со средними μ_i и общей дисперсией σ^2. Обозначим $n = \sum_{i=1}^{k} n_i$. Легко показать, что k выборочных средних $\bar{x}_i$ и объединенная сумма квадратов $S^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2$ совместно достаточны для $k+1$ параметров. Эти достаточные статистики имеют совместное распределение

$$g(\bar{x}_1, \ldots, \bar{x}_k, S^2) \propto \frac{S^{n-k-2}}{\sigma^n} \exp\left\{-\frac{1}{2\sigma^2}\sum_i\sum_j (x_{ij}-\mu_i)^2\right\}. \quad (23.99)$$

представляющее собой простое обобщение (23.96). При выводе (23.99) использована независимость $\bar{x}_i$ друг от друга и от S^2 и тот факт, что величина S^2/σ^2 распределена по закону χ^2 с $n-k$ степенями свободы. (23.99) можно также записать в форме (23.73):

$$g \propto C(\mu_i, \sigma^2)\exp\left\{\left(-\frac{1}{2}\sum_i\sum_j x_{ij}^2\right)\left(\frac{1}{\sigma^2}\right) + \sum_i\left(\sum_j x_{ij}\right)\left(\frac{\mu_i}{\sigma^2}\right)\right\}. \quad (23.100)$$

Рассмотрим теперь линейную функцию

$$\theta = A\left(\frac{1}{\sigma^2}\right) + \sum_{i=1}^{k} B_i\left(\frac{\mu_i}{\sigma^2}\right). \quad (23.101)$$

(1) Положим $A = 1$, $B_i = 0$ (для всех i). Тогда $\theta = 1/\sigma^2$, а $\psi_i = \mu_i/\sigma^2$ $(i = 1, \ldots, k)$ — мешающие параметры. Для каждой из рассмотренных в **23.30—32** четырех гипотез $H_0^{(r)}$, относящихся к параметру $1/\sigma^2$ (а следовательно, и к σ^2), существует

РНМН критерий. Эти критерии основаны на условном распределении $\sum\sum x_{ij}^2$ при фиксированном векторе $\left(\sum_j x_{1j}, \sum_j x_{2j}, \ldots, \sum_j x_{kj}\right)$, т. е. на условном распределении $S^2 = \sum_i \sum_j (x_{ij} - \bar{x}_i)^2$ (при том же условии). Точно так же, как в примере 23.14, это приводит к использованию безусловного распределения S^2 для получения РНМН критериев.

(2) Рассмотрения, полностью аналогичные приведенным в примере 23.14 (2), показывают, что, полагая $\theta_0 = 0$, мы получим РНМН критерии для $\sum_{i=1}^{k} c_i\mu_i \leqslant c_0$ и $\sum c_i\mu_i = c_0$, где c_0 — произвольная константа (см. упражнение 23.19). Как и раньше, использование этого метода, связанного с линейной формой $\sum c_i\mu_i$, не позволяет проверять «интервальные» гипотезы.

(3) При $k = 2$, $c_1 = 1$, $c_2 = -1$, $c_0 = 0$ случай (2) сводится к проверке гипотез $H_0^{(1)}$: $\mu_1 - \mu_2 \leqslant 0$, $H_0^{(2)}$: $\mu_1 - \mu_2 = 0$. Критерий для проверки гипотезы $\mu_1 - \mu_2 = 0$ был рассмотрен в примере 23.8, где было показано, что он сводится к t-критерию Стьюдента и является РНМ подобным критерием. Теперь можно показать, что он также является РНМН для $H_0^{(1)}$. Мы видим также, что двусторонний несмещенный t-критерий Стьюдента является РНМН для $H_0^{(2)}$.

Пример 23.16

Обобщим ситуацию примера 23.15, допуская, что дисперсии k нормальных распределений могут быть различными. В этом случае имеется $2k$ достаточных статистик для $2k$ параметров. Этими статистиками будут суммы и суммы квадратов $\sum_{j=1}^{n_i} x_{ij}$, $\sum_{j=1}^{n_i} x_{ij}^2$, $i = 1, 2, \ldots, k$. Положим

$$\theta = \sum_{i=1}^{k} A_i\left(\frac{1}{\sigma_i^2}\right) + \sum_{i=1}^{k} B_i\left(\frac{\mu_i}{\sigma_i^2}\right). \tag{23.102}$$

(1) Будем считать $B_i = 0$ (для всех i). Для всех четырех гипотез, относящихся к взвешенной сумме величин, обратных генеральным дисперсиям,

$$\theta = \sum_i A_i\left(\frac{1}{\sigma_i^2}\right),$$

можно получить РНМН критерии. В случае $k = 2$ имеем

$$\theta = \frac{A_1}{\sigma_1^2} + \frac{A_2}{\sigma_2^2}.$$

Если мы хотим проверять гипотезы, касающиеся отношения дисперсий σ_2^2/σ_1^2, то так же, как в случае (2) примеров 23.14, 23.15, мы должны положить $\theta = 0$. После этого РНМН критерии для $\theta = 0$, $\leqslant 0$ сводятся к РНМН критериям для

$$\frac{\sigma_2^2}{\sigma_1^2} = -\frac{A_2}{A_1}, \quad \leqslant -\frac{A_2}{A_1},$$

и мы, следовательно, получаем РНМН критерии для гипотез $H_0^{(1)}$ и $H_0^{(2)}$ об отношении дисперсий. Совместное распределение четырех достаточных статистик может быть записано в виде

$$g\left(\sum x_{1j}, \sum x_{2j}, \sum x_{1j}^2, \sum x_{2j}^2\right) \propto C(\mu_i, \sigma_i^2) \times$$
$$\times \exp\left\{-\frac{1}{2}\left(\frac{1}{\sigma_1^2}\sum x_{1j}^2 + \frac{1}{\sigma_2^2}\sum x_{2j}^2\right) + \frac{\mu_1}{\sigma_1^2}\sum x_{1j} + \frac{\mu_2}{\sigma_2^2}\sum x_{2j}\right\}. \tag{23.103}$$

Согласно **23.27** коэффициент $s(\boldsymbol{x})$ при θ в распределении (23.103), преобразованном так, чтобы θ стал одним из его параметров, является точно такой же функцией от $-\frac{1}{2}\sum x_{1j}^2$, $-\frac{1}{2}\sum x_{2j}^2$, какой является θ от $1/\sigma_1^2$, $1/\sigma_2^2$, т. е.

$$-2s(\boldsymbol{x}) = A_1 \sum x_{1j}^2 + A_2 \sum x_{2j}^2,$$

а РНМН критерии основаны на условном распределении $s(\boldsymbol{x})$ при фиксированных трех функциях от достаточных статистик, ортогональных к $s(\boldsymbol{x})$ и друг к другу, например:

$$\sum x_{1j}, \quad \sum x_{2j} \quad \text{и} \quad A_2 \sum x_{1j}^2 - A_1 \sum x_{2j}^2.$$

Это эквивалентно тому, что фиксированы

$$\bar{x}_1, \quad \bar{x}_2 \quad \text{и} \quad t = \sum (x_{1j} - \bar{x}_1)^2 - \frac{A_1}{A_2}\sum (x_{2j} - \bar{x}_2)^2,$$

так что статистика $s(\boldsymbol{x})$ эквивалентна

$$\sum (x_{1j} - \bar{x}_1)^2 + \frac{A_2}{A_1}\sum (x_{2j} - \bar{x}_2)^2$$

при фиксированном t, что, в свою очередь, эквивалентно рассмотрению распределения отношения

$$\sum (x_{1j} - \bar{x}_1)^2 / \sum (x_{2j} - \bar{x}_2)^2.$$

Таким образом, РНМН критерии для $H_0^{(1)}$, $H_0^{(2)}$ основываются на распределении выборочного отношения дисперсий (см. упражнения 23.14 и 23.17).

(2) Как следует из (23.102), мы не можем получить РНМН критериев для функций от μ_i, не зависящих от σ_i^2. В случае $k = 2$ это обстоятельство не позволяет решить данным способом задачу о двух средних.

23.34 Для лучшего понимания результатов пунктов **23.27—33** будет полезна их геометрическая интерпретация. В соответствии с (23.73) характеристическая функция $s(\boldsymbol{x})$ имеет вид

$$\varphi(u) = \mathsf{M}\{\exp(ius)\} = \frac{C(\theta)}{C(\theta + iu)},$$

и, следовательно, производящая функция семиинвариантов равна

$$\psi(u) = \log \varphi(u) = \log C(\theta) - \log C(\theta + iu). \qquad (23.104)$$

Из (23.104) следует, что r-й семиинвариант $s(\boldsymbol{x})$ равен

$$\varkappa_r = \left[\frac{\partial^r}{\partial (iu)^r} \psi(u)\right]_{u=0} = -\frac{\partial^r}{\partial \theta^r} \log C(\theta), \qquad (23.105)$$

откуда

$$\mathsf{M}(s) = \varkappa_1 = -\frac{\partial}{\partial \theta} \log C(\theta) \qquad (23.106)$$

и

$$\varkappa_r = \frac{\partial^{r-1}}{\partial \theta^{r-1}} \mathsf{M}(s), \qquad r \geqslant 2. \qquad (23.107)$$

Рассмотрим теперь производную

$$\boldsymbol{D}^q f \equiv \frac{\partial^q}{\partial \theta^q} f(x \mid \theta, \psi).$$

В силу (23.73) и (23.106)

$$\boldsymbol{D} f = \left\{ s + \frac{C'(\theta)}{C(\theta)} \right\} f = \{s - \mathsf{M}(s)\} f. \qquad (23.108)$$

Применяя правило Лейбница, из (23.108) получаем

$$\boldsymbol{D}^q f = \boldsymbol{D}^{q-1}[\{s - \mathsf{M}(s)\} f] =$$

$$= \{s - \mathsf{M}(s)\} \boldsymbol{D}^{q-1} f + \sum_{i=1}^{q-1} \binom{q-1}{i} [\boldsymbol{D}^i \{s - \mathsf{M}(s)\}][\boldsymbol{D}^{q-1} f], \qquad (23.109)$$

что в соответствии с (23.107) можно переписать в виде

$$\boldsymbol{D}^q f = \{s - \mathsf{M}(s)\} \boldsymbol{D}^{q-1} f - \sum_{i=1}^{q-1} \binom{q-1}{i} \varkappa_{i+1} \boldsymbol{D}^{q-1-i} f. \qquad (23.110)$$

23.35 Рассмотрим теперь любую критическую область w размера α. Ее функция мощности определяется формулой (23.80), и мы можем иначе записать ее в виде интеграла по выборочному пространству достаточных статистик (s, t):

$$P(w\,|\,\theta) = \int_w f\,ds\,dt, \tag{23.111}$$

где f обозначает совместную плотность (s, t), имеющую, как было показано, вид (23.73). Производные функции мощности (23.111) выражаются следующим образом:

$$P^{(q)}(w\,|\,\theta) = \int_w \boldsymbol{D}^q f\,ds\,dt, \tag{23.112}$$

так как в (23.111) можно дифференцировать под знаком интеграла. Используя (23.108) и (23.110), получаем из (23.111)

$$P'(w\,|\,\theta) = \int_w \{s - \mathbf{M}(s)\} f\,ds\,dt = \operatorname{cov}\{s,\ c(w)\} \tag{23.113}$$

и

$$P^{(q)}(w\,|\,\theta) = \int_w \{s - \mathbf{M}(s)\}\,\boldsymbol{D}^{q-1} f\,ds\,dt - \sum_{i=1}^{q-1} \binom{q-1}{i} \varkappa_{i+1} P^{(q-1-i)}(w\,|\,\theta),$$
$$q \geqslant 2. \tag{23.114}$$

Полученное рекуррентное соотношение позволяет нам выразить значение любой производной через низшие производные. В частности, из (23.114) получаем

$$P''(w\,|\,\theta) = \operatorname{cov}\{[s - \mathbf{M}(s)]^2,\ c(w)\}, \tag{23.115}$$

$$\left.\begin{aligned} P'''(w\,|\,\theta) &= \operatorname{cov}\{[s - \mathbf{M}(s)]^3,\ c(w)\} - 3\varkappa_2 P'(w\,|\,\theta), \\ P^{\mathrm{IV}}(w\,|\,\theta) &= \operatorname{cov}\{[s - \mathbf{M}(s)]^4,\ c(w)\} - \\ &\quad - 6\varkappa_2 P''(w\,|\,\theta) - 4\varkappa_3 P'(w\,|\,\theta). \end{aligned}\right\} \tag{23.116}$$

Из (23.113) и (23.115) следует, что первые две производные представляют собой просто ковариации между $c(w)$ и s и между $c(w)$ и квадратом отклонения s от своего среднего. Третья и четвертая производные, даваемые формулами (23.116), а также высшие производные являются более сложными функциями от ковариаций и семиинвариантов s.

23.36 Теперь мы можем перейти к геометрической интерпретации некоторых результатов, полученных в **23.27—33.** Чтобы максимизировать мощность, мы должны в соответствии с (23.113) и (23.114) выбрать w так, чтобы максимизировать (для всех допустимых альтернатив) ковариацию между $c(w)$ и s или

некоторой функцией от $s - \mathbf{M}(s)$. Очевидно, что в $(r+1)$-мерном пространстве достаточных статистик (s, t) это можно сделать, ограничившись подпространством, ортогональным к r координатам, соответствующим компонентам t, т. е. ограничившись условным распределением s при фиксированном t.

В случае проверки гипотезы $\theta = \theta_0$ против $\theta > \theta_0$ мы получаем максимум $P(w|\theta)$ для всех $\theta > \theta_0$, максимизируя $P'(w|\theta)$, т. е. максимизируя $\operatorname{cov}(s, c(w))$ для всех $\theta > \theta_0$. Легко видеть, что эта цель достигается, если w состоит из 100α процентов *наибольших* значений распределения s при фиксированном t. Аналогично, в случае проверки $\theta = \theta_0$ против $\theta < \theta_0$ мы максимизируем P, *минимизируя* P', что может быть достигнуто, если w состоит из 100α процентов наименьших значений распределения s при фиксированном t. Поскольку $P'(w|\theta)$ не меняет знака, односторонние критерии несмещены.

В случае двусторонней гипотезы $H_0^{(2)}$ из **23.31** мы должны согласно (23.81) и (23.115) максимизировать $P''(w|\theta)$, т. е. $\operatorname{cov}\{[s - \mathbf{M}(s)]^2, c(w)\}$. Те же рассуждения, что и в одностороннем случае, показывают, что нужно выбрать область w так, чтобы она содержала 100α процентов наибольших значений $\{s - \mathbf{M}(s)\}^2$, так что мы получим двустороннюю критическую область, которая будет симметричной только в том случае, если распределение s при фиксированном t симметрично. Отсюда следует, что границы РНМН критической области находятся на равном расстоянии от $\mathbf{M}(s|t)$.

Вспомогательные статистики: принцип условности

23.37 Мы знаем, что всегда существует минимальная достаточная система (T_r, T_s) из $r+s$ $(r \geqslant 1,\ s \geqslant 0)$ статистик для $k+l$ $(k \geqslant 1,\ l \geqslant 0)$ параметров (θ_k, θ_l). Предположим теперь, что распределение подсистемы T_s не зависит от θ_k. (Это возможно только в том случае, если распределение (T_r, T_s) неполно, — см. упражнение 23.7.) Тогда функция правдоподобия факторизуется следующим образом:

$$L(\bar{x}|\,\theta_k,\ \theta_l) = g(T_r,\ T_s\,|\theta_k,\ \theta_l)\,h(\boldsymbol{x}) = g_1(T_r\,|\,T_s,\ \theta_k,\ \theta_l)\,g_2(T_s|\ \theta_l)\,h(\boldsymbol{x}). \tag{23.117}$$

Это есть соотношение (21.95), записанное в других обозначениях.

Фишер (например, 1956) называет T_s *вспомогательной статистикой* *), а Бартлетт (например, 1939) называет условную статистику $(T_r|T_s)$ *квазидостаточной* для θ_k из-за сходства

*) Английский термин: ancillary statistic. (*Прим. ред.*)

(23.117) при известном θ_l с факторизацией (17.84), характеризующей достаточную статистику.

Фишер предложил использовать для статистических выводов вообще и для проверки гипотез в частности *принцип условности*, заключающийся в следующем: если распределение T_s не зависит от θ_k, как в (23.117), то для получения выводов о θ_k нам нужно только условное распределение $T_r|T_s$. Если T_s достаточна для θ_l при известном θ_k, то из (23.117) сразу получаем представление

$$L(\boldsymbol{x}|\,\theta_k,\,\theta_l) = g_1(T_r|\,T_s,\,\theta_k)\,g_2(T_s|\,\theta_l)\,h(\boldsymbol{x}), \qquad (23.118)$$

показывающее, что распределения $(T_r|T_s)$ и T_s зависят каждое от своего параметра и каждое достаточно для своего параметра. В этом случае, без сомнения, в соответствии с общим принципом пункта **23.3** при проверке гипотез о θ_k можно ограничиться только функциями от $(T_r|T_s)$.

Однако основной вопрос заключается в том, можем ли мы ограничиться условной статистикой, когда T_s не достаточна для θ_l. В самом деле, Уэлч (1939) привел пример простой гипотезы о среднем значении прямоугольного распределения с известным размахом, показывающий, что условный критерий, основанный на $(T_r|T_s)$, может оказаться равномерно менее мощным, чем другой (безусловный) критерий.

С ответом на этот вопрос связаны далеко идущие выводы, так как А. Бирнбаум (1962) показал, что принцип условности влечет принцип правдоподобия (и, конечно, следует из последнего), который утверждает (см. **18.32**), что любой статистический вывод из наблюдений требует только знания ФП. В частности, из этого принципа следует, что знание деталей выборочной процедуры, с помощью которой получаются наблюдения (и ФП), ничего не дает для последующего статистического вывода.

Многие (возможно, большинство) из статистиков сочтут интуитивно неприемлемым исключать выборочное пространство из рассмотрения при осуществлении статистических выводов. Придерживаясь этой точки зрения, они должны отказаться от принципа правдоподобия, а вместе с ним, автоматически, и от принципа условности.

Пример 23.17

В примере 17.17 было показано, что пара статистик $(\bar{x}, s^2)$ совместно достаточна для параметров (μ, σ^2) нормального распределения и что распределение s^2 не зависит от μ. Таким образом, мы имеем соотношение

$$L(x|\,\mu,\,\sigma^2) = g_1[(\bar{x}|\,s^2)|\,\mu,\,\sigma^2]\,g_2(s^2|\,\sigma^2)\,h(x),$$

являющееся частным случаем (23.117) при $k = l = r = s = 1$. Принцип условности утверждает, что для проверки гипотез о μ следует пользоваться условной статистикой $x|s^2$ (так получилось, что $\bar{x}$ в данном случае не зависит от s^2, но это упрощение никак не связано с аргументацией в общем случае). Однако s^2 не является достаточной статистикой для мешающего параметра σ^2, так что распределение $\bar{x}|s^2$ зависит от σ^2. Если априорное распределение σ^2 неизвестно, то достичь какого-то успеха можно, лишь более или менее произвольно усредняя по σ^2. Если мы согласны использовать фидуциальное распределение σ^2 и интегрировать по нему, то мы возвратимся к ситуации пункта **21.10**, где мы получили результат, совпадающий с результатом, полученным при максимизации мощности в примерах 23.7 и 23.14 и состоящим в использовании t-распределения Стьюдента.

Другой принцип условного критерия

23.38 Другой принцип построения критериев (см. Д. Кокс (1958a)) заключается в использовании $(T_r|T_s)$ всегда, когда статистика T_s достаточна для θ_l при известном θ_k, независимо от того, зависит ее распределение от θ_k или нет. В этом случае мы имеем

$$L(\boldsymbol{x}|\,\theta_k,\,\theta_l) = g_1(T_r\,|\,T_s,\,\theta_k)\,g_2(T_s|\,\theta_k,\,\theta_l)\,h(\boldsymbol{x}), \qquad (23.119)$$

так что распределение условной статистики не зависит от мешающего параметра θ_l. При таком подходе мы также не можем предполагать, что полученный критерий будет в каком-то смысле оптимальным. Соотношение (23.119) есть просто (21.96) в других обозначениях.

Обоснование условных критериев

23.39 Из **23.30—32** вытекает, что если распределение достаточных статистик (T_r, T_s) имеет экспоненциальную форму (23.73), то использование условного распределения T_r при фиксированном T_s приводит к РНМН критериям (T_r и T_s являются просто новыми обозначениями для $s(\boldsymbol{x})$ и $\boldsymbol{t}(\boldsymbol{x})$). Если же распределение достаточных статистик не принадлежит к типу (23.73) (например, если оно сосредоточено на интервале, концы которого зависят от параметра), то приведенное обоснование становится недействительным. Однако Линдли (1958b) показал, что можно обосновать использование условной статистики $T_r|T_s$, предполагая только, что распределение $g_2(T_s|\theta_k, \theta_l)$ статистики T_s ограниченно полно при гипотезе H_0 и что T_s достаточна для θ_l. Действительно, тогда в силу **23.19** любая критическая область размера α, подобная относительно θ_l, содержит долю α

каждой поверхности, соответствующей постоянному значению T_s. Таким образом, любой подобный критерий для H_0 будет условным критерием, основанным на $T_r|T_s$, а любой оптимальный условный критерий будет оптимальным подобным критерием.

Контрпример Уэлча (1939), приводимый в упражнении 23.23, не попадает ни под одно из наших обоснований использования условных статистик критериев, так как имеющаяся в этом примере двухкомпонентная минимальная достаточная статистика для единственного параметра неполна, и поэтому результат упражнения 23.7 не противоречит существованию вспомогательной статистики (выборочного размаха R).

УПРАЖНЕНИЯ

23.1 Показать, что для случая выборок объема n из нормального распределения со средним θ и дисперсией σ^2 не существует симметричной области, подобной относительно θ и σ^2 при $n \leqslant 2$, но что такая область существует при $n \geqslant 3$.

(Феллер, 1938.)

23.2 Аналогично тому, как было сделано в примере 23.1, показать, что в случае выборки из n наблюдений, i-е из которых имеет распределение

$$\frac{1}{\Gamma(\theta_i)} e^{-x_i} x_i^{\theta_i - 1}\, dx_i, \qquad 0 \leqslant x_i \leqslant \infty, \quad \theta_i > 0,$$

не существует подобной области размера α при $0 < \alpha < 1$.

(Феллер, 1938.)

23.3 Пусть $L(x|\theta)$ — функция правдоподобия и $\mathbf{M}\left(\frac{\partial \log L}{\partial \theta}\right) = 0$. Показать, что если распределение статистики z не зависит от θ, то $\operatorname{cov}\left(z, \frac{\partial \log L}{\partial \theta}\right) = 0$. Вывести отсюда, что если не существует статистики, некоррелированной с $\frac{\partial \log L}{\partial \theta}$, то не существует областей, подобных относительно θ.

(Нейман, 1938а.)

23.4 Используя х. ф. z, показать, что верны также утверждения, обратные утверждению упражнения 23.3 и его следствию.

Таким образом, это и предыдущее упражнение вместе устанавливают, что условие $\operatorname{cov}\left(z, \frac{\partial \log L}{\partial \theta}\right) = 0$ необходимо и достаточно для выполнения соотношения $\operatorname{cov}\left(e^{iuz}, \frac{\partial \log L}{\partial \theta}\right) = 0$, где u — произвольная переменная.

(Нейман, 1938а.)

23.5 Показать, что семейство распределений Коши

$$dF = \frac{dx}{\{\pi\theta^{1/2}(1 + (x^2/\theta))\}}, \qquad -\infty \leqslant x \leqslant \infty,$$

неполно.

(Леман и Шеффе, 1950.)

23.6 Показать, что если статистика z не зависит от достаточной для θ статистики t, то распределение z не зависит от θ.

23.7 Пусть в упражнении 23.6 $H_1(z)$ обозначает ф. р. z, $H_2(z|t)$ — условную ф. р. при фиксированном t и $g(t|\theta)$ — плотность t. Показать, что

$$\int \{H_1(z) - H_2(z|t)\}\, g(t|\theta)\, dt = 0$$

для всех θ. Вывести отсюда, что если t — *полная* достаточная статистика для θ, то справедлив результат, обратный к полученному в упражнении 23.6, именно, если распределение z не зависит от θ, то z не зависит от t.

(Д. Басу, 1955.)

23.8 Использовать результат упражнения 23.7 для прямого доказательства того, что в случае выборок из одномерного нормального распределения

(а) любой выборочный момент относительно выборочного среднего $\bar{x}$ не зависит от $\bar{x}$;

(б) квадратичная форма $\boldsymbol{x'Ax}$ не зависит от $\bar{x}$ тогда и только тогда, когда сумма элементов каждой строки матрицы $\boldsymbol{A}$ равна нулю (см. **15.15**);

(в) выборочный размах не зависит от $\bar{x}$;

(г) $(x_{(n)} - \bar{x})/(x_{(n)} - x_{(1)})$ не зависит как от $\bar{x}$, так и от выборочной дисперсии s^2.

(Хогг и Крэйг, 1956.)

23.9 Используя упражнение 23.7, показать, что

(а) в случае выборок из двумерного нормального распределения с $\rho = 0$ выборочный коэффициент корреляции не зависит от выборочных среднего и дисперсии (см. **16.28**);

(б) в случае независимых выборок из двух одномерных нормальных совокупностей с одной и той же дисперсией σ^2 статистика

$$F = \frac{\sum_j (x_{1j} - \bar{x}_1)^2/(n_1 - 1)}{\sum_j (x_{2j} - \bar{x}_2)^2/(n_2 - 1)}$$

не зависит от системы из трех совместно достаточных статистик

$$\bar{x}_1,\ \bar{x}_2,\ \sum_j (x_{1j} - \bar{x}_1)^2 + \sum_j (x_{2j} - \bar{x}_2)^2$$

и, следовательно, от статистики

$$t^2 = \frac{(\bar{x}_1 - \bar{x}_2)^2}{\sum (x_{1j} - \bar{x}_1)^2 + \sum (x_{2j} - \bar{x}_2)^2} \left\{ \frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2} \right\},$$

являющейся функцией от достаточных статистик. Это утверждение верно независимо от того, равны или нет генеральные средние.

(Хогг и Крэйг, 1956.)

23.10 Показать, что в случае выборки объема n из распределения

$$dF = \exp\{-(x - \theta)\}\, dx, \qquad \theta \leqslant x \leqslant \infty,$$

распределение $x_{(1)}$ не зависит от

$$z = \sum_{i=1}^{r} (x_{(i)} - x_{(1)}) + (n - r)(x_{(r)} - x_{(1)}), \qquad r \leqslant n.$$

(Эпстейн и Собел, 1954.)

23.11 Показать, что в случае биномиального распределения с параметром π выборочная частота p служит минимальной достаточной статистикой для π.

(Леман и Шеффе, 1950.)

23.12 Показать, что в случае прямоугольного распределения

$$dF = dx, \qquad \theta - \frac{1}{2} \leqslant x \leqslant \theta + \frac{1}{2},$$

пара статистик $(x_{(1)}, x_{(n)})$ служит минимальной достаточной системой для θ.

(Леман и Шеффе, 1950.)

23.13 Методом пункта **23.20** показать, что в случае нормального распределения с дисперсией σ^2 и неизвестным средним μ РНМ подобный критерий для H_0: $\sigma^2 = \sigma_0^2$ против H_1: $\sigma^2 = \sigma_1^2$ принимает вид

$$\sum (x - \bar{x})^2 \geqslant a_\alpha, \qquad \text{если} \qquad \sigma_1^2 > \sigma_0^2,$$
$$\sum (x - \bar{x})^2 \leqslant b_\alpha, \qquad \text{если} \qquad \sigma_1^2 < \sigma_0^2.$$

23.14 Два нормальных распределения имеют неизвестные средние и дисперсии σ^2, $\theta\sigma^2$. Для двух независимых выборок объемов n_1, n_2 из этих распределений методом пункта **23.20** показать, что РНМ подобный критерий для H_0: $\theta = 1$ против H_1: $\theta = \theta_1$ принимает вид

$$s_1^2/s_2^2 \geqslant a_\alpha, \qquad \text{если} \qquad \theta_1 > 1,$$
$$s_1^2/s_2^2 \leqslant b_\alpha, \qquad \text{если} \qquad \theta_1 < 1,$$

где s_1^2, s_2^2 — выборочные дисперсии.

(Хартер (1963) показал, что критерий, основанный на отношении выборочных размахов, имеет почти такую же мощность, как и РНМ подобный критерий.)

23.15 Две независимые выборки, каждая объема n, взяты из распределений

$$dF = \exp\left(-\frac{x}{\theta_1}\right)\frac{dx}{\theta_1}, \qquad dG = \exp(-y\theta_2)\,\theta_2\,dy,$$
$$\theta_1, \theta_2 > 0, \qquad 0 \leqslant x, y \leqslant \infty.$$

Показать, что $t = \left(\sum x, \sum y\right) = (X, Y)$ есть минимальная достаточная система для (θ_1, θ_2), остающаяся таковой также при выполнении гипотезы H_0: $\theta_1 = \theta_2 = \theta$. Рассматривая функцию $XY - \mathbf{M}(XY)$, показать, что распределение t не является ограниченно полным при гипотезе H_0, так что не все подобные области удовлетворяют (23.13). Показать, что статистика XY в этом случае распределена независимо от θ, так что подобные критические области для проверки H_0 можно строить с помощью распределения XY.

(Ватсон, 1957а.)

23.16 Используя (23.98), показать, что в примере 23.14 имеется РНМН критерий для гипотезы о том, что параметрическая точка (μ, σ) лежит между двумя параболами

$$\mu = \mu_0 + c_1\sigma^2, \qquad \mu = \mu_0 + c_2\sigma^2,$$

касательными одна к другой в точке $(\mu_0, 0)$.

(Леман и Шеффе, 1955.)

23.17 Показать, что в упражнении 23.14 критическая область

$$s_1^2/s_2^2 \geqslant a_{\alpha/2}, \qquad \leqslant b_{\alpha/2}$$

для двусторонней альтернативы H_1: $\theta \neq 1$ смещена, если $n_1 \neq n_2$. Пользуясь методом примера 23.12, показать, что несмещенная критическая область

$$t = s_1^2/s_2^2 \geqslant a_{1-\alpha_1}, \qquad \leqslant b_{\alpha_2}, \qquad \alpha_1 + \alpha_2 = \alpha,$$

определяется условием (см. (23.56))

$$V_{\alpha_2} f(V_{\alpha_2}) = V_{1-\alpha_1} f(V_{1-\alpha_1}),$$

где f — плотность отношения дисперсий t, а V_α — ее 100α-процентная точка. Показать также, что функция мощности несмещенного критерия монотонно возрастает при $\theta > 1$ и монотонно убывает при $\theta < 1$.

(Рамачандран (1958) приводит значения $V_{1-\alpha_1}$, V_{α_2} для $\alpha = 0{,}05$, $n_1 - 1$ и $n_2 - 1 = 2(1)4(2)12(4)24, 30, 40, 60$.)

23.18 Показать, что в упражнении 23.17 несмещенный доверительный интервал $\left(\dfrac{t}{V_{\alpha_2}}, \dfrac{t}{V_{1-\alpha_1}}\right)$ для θ минимизирует математическое ожидание $(\log U - \log L)$ в классе доверительных интервалов (L, U), основанных на «хвостах» распределения t.

(Шеффе, 1942b.)

23.19 Используя **23.27**, показать, что в примере 23.15 РНМН критерии для $\sum_i c_i \mu_i$ основаны на распределении статистики

$$t = \sum c_i (\bar{x}_i - \mu_i) \Big/ \left\{ \frac{S^2}{n-k} \sum \frac{c_i^2}{n_i} \right\}^{1/2},$$

имеющей t-распределение Стьюдента с $n - k$ степенями свободы.

23.20 Показать, что в примере 23.16 существует РНМН критерий для гипотезы

$$\frac{\mu_i}{\mu_j} = a \frac{\sigma_i^2}{\sigma_j^2}, \qquad i \neq j, \qquad \mu_j \neq 0.$$

23.21 Для случая независимых выборок из двух пуассоновских распределений с параметрами μ_1, μ_2 показать, что для всех четырех гипотез пунктов **23.30—32**, касающихся μ_1/μ_2, существуют РНМН критерии и что критерий для $\mu_1/\mu_2 = 1$ состоит в проверке того, что сумма наблюдений распределена между двумя выборками по биномиальному закону с равными вероятностями.

(Леман и Шеффе, 1955.)

23.22 Для случая независимых биномиальных распределений с параметрами θ_1, θ_2 найти РНМН критерии для всех четырех гипотез пунктов **23.30—32**, касающихся «отношения шансов» $\left(\dfrac{\theta_1}{1-\theta_1}\right) \Big/ \left(\dfrac{\theta_2}{1-\theta_2}\right)$, а также РНМН критерии для $\theta_1 = \theta_2$, $\theta_1 \leqslant \theta_2$.

(Леман и Шеффе, 1955.)

23.23 Для прямоугольного распределения

$$dF = dx, \qquad \theta - \frac{1}{2} \leqslant x \leqslant \theta + \frac{1}{2},$$

условное распределение средины размаха M при фиксированном размахе R и маргинальное распределение M получены в упражнениях 14.12, 14.13 и 4.16. Показать, что при проверке гипотезы H_0: $\theta = \theta_0$ против двусторонней альтернативы H_1: $\theta \neq \theta_0$ критерий с равными «хвостами», основанный на распределении M при фиксированном R, после интегрирования по R дает критерий равномерно менее мощный, чем критерий с равными «хвостами», основанный на маргинальном распределении M. Для удобства использовать $\alpha = 0{,}08$.

(Уэлч, 1939.)

23.24 Показать, что в примере 23.9 РНМН критерий для проверки H_0: $\sigma = \sigma_0$ против H_1: $\sigma \neq \sigma_0$ имеет вид

$$\sum_{i=1}^{n} x_i \geqslant a_{\alpha_1}, \qquad \leqslant b_{\alpha_2}.$$

(Леман, 1947.)

23.25 Показать, что в случае распределения примера 23.9 РНМ подобный критерий для проверки H_0: $\theta = \theta_0$ против H_1: $\theta \neq \theta_0$ имеет вид

$$\frac{x_{(1)} - \theta_0}{\bar{x}} < 0, \qquad \geqslant c_\alpha.$$

(Леман, 1947.)

23.26 Показать, что в случае прямоугольного распределения

$$dF = dx/\theta, \qquad \mu \leqslant x \leqslant \mu + \theta,$$

РНМ подобный критерий для проверки H_0: $\mu = \mu_0$ против H_1: $\mu \neq \mu_0$ имеет вид

$$\frac{x_{(1)} - \mu_0}{x_{(n)} - x_{(1)}} < 0, \qquad \geqslant c_\alpha.$$

Ср. со случаем простой гипотезы при $\theta = 1$, когда, как было показано в упражнении 22.8, РНМ критерия не существует.

(Леман, 1947.)

23.27 Пусть $x_1, \ldots, x_n$ — независимые наблюдения из распределения

$$dF = \frac{1}{\theta^p \Gamma(p)} \exp(-x/\theta)\, x^{p-1}\, dx, \qquad p > 0, \qquad 0 \leqslant x \leqslant \infty.$$

Используя упражнения 23.6 и 23.7, показать, что для того, чтобы статистика $h(x_1, \ldots, x_n)$ не зависела от $S = \sum_{i=1}^{n} x_i$, необходимо и достаточно, чтобы $h(x_1, \ldots, x_n)$ была однородной функцией от x степени 0 (см. ссылки к упражнению 15.22).

23.28 Используя (23.113) и (23.114), показать, что если первая отличная от нуля производная функции мощности имеет порядок m, то

$$P^{(m)}(w \mid \theta) = \operatorname{cov}\{[s - \mathbf{M}(s)]^m,\ c(w)\}$$

и

$$\frac{\{P^{(m)}(w \mid \theta)\}^2}{\mu_{2m}} \leqslant \frac{1}{4},$$

где μ_r есть r-й центральный момент s. В частности,

$$|P'(w|\theta)| \leqslant \frac{1}{2}\mu_2^{1/2}.$$

23.29 Используя **23.35**, показать, что w является подобной областью для некоторой гипотезы с мешающим параметром θ тогда и только тогда, когда

$$\operatorname{cov}\{s, c(w)\} = 0$$

тождественно по θ. См. упражнения 23.3, 23.4.

23.30 Обобщая доказательство, приведенное в конце примера 23.7, показать, что для любого распределения вида

$$dF = f\left(\frac{x-\mu}{\sigma}\right)\frac{dx}{\sigma},$$

обладающего полной достаточной статистикой для σ при известном μ, не может существовать подобной области для H_0: $\mu = \mu_0$ против H_1: $\mu = \mu_1$, мощность которой не зависит от σ.

23.31 Для случая нормального распределения со средним и дисперсией, равными θ, показать, что при объеме выборки $n = 1$ обе статистики x и x^2 являются достаточными для θ, причем только x^2 будет минимальной. Отсюда следует, что одномерная достаточность не влечет минимальной достаточности. Упражнение 18.13 иллюстрирует этот же факт в двумерном случае.

ГЛАВА 24

КРИТЕРИИ ОТНОШЕНИЯ ПРАВДОПОДОБИЯ И ОБЩАЯ ЛИНЕЙНАЯ ГИПОТЕЗА

24.1 Разобранный в главе 18 метод МП является конструктивным методом для получения оценок, которые (при определенных условиях) обладают желаемыми свойствами. К нему близок предложенный Нейманом и Пирсоном (1928) метод отношения правдоподобия (ОП), используемый для проверки гипотез. Он играет такую же роль в теории проверки гипотез, как метод МП в теории оценивания.

Как и прежде, у нас имеется ФП

$$L(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta),$$

где $\theta = (\theta_r, \theta_s)$ — вектор $r + s = k$ параметров ($r \geqslant 1$, $s \geqslant 0$), а x также может быть вектором. Мы хотим проверить гипотезу

$$H_0\colon\ \theta_r = \theta_{r0}, \tag{24.1}$$

которая при $s \neq 0$ является сложной, против

$$H_1\colon\ \theta_r \neq \theta_{r0}.$$

Известно, что в этой задаче, вообще говоря, нет РНМ критерия, но может существовать РНМН критерий (см. **23.31**).

По методу ОП вначале требуется найти МП-оценки вектора (θ_r, θ_s), дающие безусловный максимум ФП

$$L(x|\hat{\theta}_r, \hat{\theta}_s), \tag{24.2}$$

и, кроме того, МП-оценки для θ_s в предположении, что имеет место H_0 *), при которых достигается условный максимум ФП

$$L(x|\theta_{r0}, \hat{\hat{\theta}}_s). \tag{24.3}$$

В (24.3) использовано обозначение $\hat{\hat{\theta}}_s$, указывающее на то, что в общем случае эти оценки не совпадают с $\hat{\theta}_s$ из (24.2).

*) Когда $s = 0$, H_0 является простой гипотезой, и не нужно искать максимум, так как тогда значение L определено однозначно.

Рассмотрим теперь отношение правдоподобия *)

$$l = \frac{L(x \mid \theta_{r0}, \hat{\hat{\theta}}_s)}{L(x \mid \hat{\theta}_r, \hat{\theta}_s)}. \tag{24.4}$$

Поскольку (24.4) является отношением условного максимума ФП к безусловному, то, очевидно,

$$0 \leqslant l \leqslant 1. \tag{24.5}$$

Интуитивно понятно, что l является разумной статистикой критерия для проверки H_0. Действительно, она представляет собой максимум правдоподобия при гипотезе H_0, отнесенный к своей наибольшей возможной величине, и большое значение l указывает на то, что разумно принять H_0. Критическая область этого критерия имеет, следовательно, вид

$$l \leqslant c_\alpha, \tag{24.6}$$

где c_α определяется по распределению $g(l)$ статистики l так, чтобы получить критерий размера α, а именно:

$$\int_0^{c_\alpha} g(l)\, dl = \alpha. \tag{24.7}$$

24.2 Для того чтобы метод ОП мог быть полезным при построении подобных критериев, т. е. критериев, основанных на подобных критических областях, распределение статистики l не должно зависеть от мешающих параметров. Именно так обстоит дело в большинстве статистических задач. Следующие два примера иллюстрируют предложенный метод. В одном из них он приводит к подобному критерию, а в другом нет.

Пример 24.1

Для нормального распределения

$$dF(x) = (2\pi\sigma^2)^{-1/2} \exp\left\{ -\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \right\} dx$$

проверяется гипотеза

$$H_0\colon \mu = \mu_0.$$

Здесь

$$L(x \mid \mu, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left\{ -\frac{1}{2} \sum \left(\frac{x-\mu}{\sigma}\right)^2 \right\}.$$

*) Это отношение обычно обозначают λ, а статистику ОП иногда называют «ламбда-статистикой». Здесь мы используем латинскую букву, так как греческими буквами, как ранее было условлено, обозначаются параметры.

Используя пример 18.11, получаем безусловные МП-оценки

$$\hat{\mu} = \bar{x}, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum (x - \bar{x})^2 = s^2,$$

так что

$$L(x|\,\hat{\mu},\ \hat{\sigma}^2) = (2\pi s^2)^{-n/2} \exp(-n/2). \tag{24.8}$$

Когда справедлива гипотеза H_0, МП-оценкой служит (см. пример 18.8)

$$\hat{\hat{\sigma}}^2 = \frac{1}{n}\sum (x - \mu_0)^2 = s^2 + (\bar{x} - \mu_0)^2,$$

и поэтому

$$L(x|\,\mu_0,\ \hat{\hat{\sigma}}^2) = [2\pi\{s^2 + (\bar{x} - \mu_0)^2\}]^{-n/2} \exp(-n/2). \tag{24.9}$$

Из (24.4), (24.8) и (24.9) находим

$$l = \left\{ \frac{s^2}{s^2 + (\bar{x} - \mu_0)^2} \right\}^{n/2}, \qquad \text{или} \qquad l^{2/n} = \frac{1}{1 + \dfrac{t^2}{n-1}},$$

где t имеет t-распределение Стьюдента с $n - 1$ степенями свободы. Отсюда видно, что l является монотонно убывающей функцией от t^2. Следовательно, вместо распределения величины l можно использовать известное распределение статистики t^2, отвергая 100α процентов наибольших значений t^2, что соответствует 100α процентам наименьших значений l. Таким образом, получен критерий, основанный на распределении t-статистики Стьюдента, у которого половина критической области состоит из крайних положительных, а половина — из крайних отрицательных значений величины t. Это вполне разумный критерий, и в примере 23.14 мы видели, что он является РНМН критерием для гипотезы H_0.

Пример 24.2

Рассмотрим снова проблему двух средних, подробно обсуждавшуюся в главах 21 и 23. По выборкам объемов n_1, n_2 из нормальных совокупностей с параметрами (μ_1, σ_1^2), (μ_2, σ_2^2) нужно проверить гипотезу H_0: $\mu_1 = \mu_2$, которую можно представить (см. 23.2) в виде H_0: $\theta \equiv \mu_1 - \mu_2 = 0$. Обозначим μ общее неизвестное значение среднего. Имеем

$$L(x|\,\mu_1,\ \mu_2,\ \sigma_1^2,\ \sigma_2^2) =$$

$$= (2\pi)^{-(n_1+n_2)/2}\, \sigma_1^{-n_1} \sigma_2^{-n_2} \exp\left\{ -\frac{1}{2}\left(\sum_{j=1}^{n_1} \frac{(x_{1j} - \mu_1)^2}{\sigma_1^2} + \sum_{j=1}^{n_2} \frac{(x_{2j} - \mu_2)^2}{\sigma_2^2} \right) \right\}.$$

Безусловными МП-оценками служат

$$\hat{\mu}_1 = \bar{x}_1, \quad \hat{\mu}_2 = \bar{x}_2, \quad \hat{\sigma}_1^2 = s_1^2, \quad \hat{\sigma}_2^2 = s_2^2,$$

так что

$$L\left(x \mid \hat{\mu}_1,\ \hat{\mu}_2,\ \hat{\sigma}_1^2,\ \hat{\sigma}_2^2\right) = (2\pi)^{-(n_1+n_2)/2}\, s_1^{-n_1} s_2^{-n_2} \exp\left\{-\frac{1}{2}(n_1+n_2)\right\}.$$

Когда выполняется гипотеза H_0, то МП-оценки находятся из системы уравнений

$$\left.\begin{aligned} &\frac{n_1(\bar{x}_1-\mu)}{\sigma_1^2} + \frac{n_2(\bar{x}_2-\mu)}{\sigma_2^2} = 0, \\ &\sigma_1^2 = \frac{1}{n_1}\sum_{j=1}^{n_1}(x_{1j}-\mu)^2 = s_1^2 + (\bar{x}_1-\mu)^2, \\ &\sigma_2^2 = \frac{1}{n_2}\sum_{j=1}^{n_2}(x_{2j}-\mu)^2 = s_2^2 + (\bar{x}_2-\mu)^2. \end{aligned}\right\} \qquad (24.10)$$

Подставляя решение системы (24.10) в ФП, получаем

$$L\left(x \mid \hat{\hat{\mu}},\ \hat{\hat{\sigma}}_1^2,\ \hat{\hat{\sigma}}_2^2\right) = (2\pi)^{-(n_1+n_2)/2}\, \hat{\hat{\sigma}}_1^{-n_1} \hat{\hat{\sigma}}_2^{-n_2} \exp\left\{-\frac{1}{2}(n_1+n_2)\right\},$$

и отношение правдоподобия равно

$$l = \left(\frac{s_1}{\hat{\hat{\sigma}}_1}\right)^{n_1}\left(\frac{s_2}{\hat{\hat{\sigma}}_2}\right)^{n_2} = \left\{\frac{s_1^2}{s_1^2+(\bar{x}_1-\hat{\hat{\mu}})^2}\right\}^{n_1/2}\left\{\frac{s_2^2}{s_2^2+(\bar{x}_2-\hat{\hat{\mu}})^2}\right\}^{n_2/2}. \qquad (24.11)$$

Чтобы использовать (24.11), нужно определить $\hat{\hat{\mu}}$. Из (24.10) видно, что $\hat{\hat{\mu}}$ является решением кубического по μ уравнения, коэффициенты которого зависят от n_i, от сумм наблюдений и от сумм их квадратов по каждой выборке. Мы не можем выписать $\hat{\hat{\mu}}$ в виде явной функции, хотя в каждой конкретной задаче уравнение можно решить численно. Во всяком случае распределение получаемого решения зависит от отношения σ_1^2/σ_2^2, потому что $\hat{\hat{\mu}}$ есть функция от s_1^2 и s_2^2. Следовательно, статистика l имеет вид

$$l = g\left(s_1^2,\ s_2^2\right) h\left(s_1^2,\ s_2^2\right).$$

Таким образом, метод ОП не дает в этом случае подобного критерия.

24.3 Если (как в примере 24.1) оказывается, что статистика ОП является взаимно однозначной функцией некоторой статистики, распределение которой либо известно точно (как в

упомянутом примере), либо может быть найдено, то нетрудно построить критерий для проверки гипотезы H_0, хотя вопрос о том, какими желательными свойствами обладают критерии ОП, еще ожидает решения. Однако часто оказывается, что метод ОП не так удобен. Это происходит тогда, когда статистика ОП представляет собой более или менее сложную функцию от наблюдений, точное распределение которой не может быть получено (как в примере 24.2). В таком случае приходится аппроксимировать ее распределение.

Поскольку статистика l распределена на интервале (0, 1), то для любой фиксированной константы $c > 0$ величина $w = -2c \log l$ принимает значения на интервале $(0, \infty)$. Естественно поэтому аппроксимировать распределение последней распределением χ^2, сосредоточенным также на интервале $(0, \infty)$, подбирая константу так, чтобы сделать приближение как можно более точным. Причина использования такой аппроксимации состоит в том, что с ростом n распределение величины $-2 \log l$ при гипотезе H_0 (как будет доказано в **24.7**) стремится к распределению χ^2 с r степенями свободы. На самом деле мы сможем найти асимптотическое распределение для $-2 \log l$ и при гипотезе H_1, но для этого нам необходимо ввести некоторое обобщение χ^2-распределения.

Нецентральное χ^2-распределение

24.4 В **16.2—3** мы видели, что сумма квадратов n независимых нормированных нормальных величин имеет распределение χ^2 с n степенями свободы (16.1), а ее х. ф. дается формулой (16.3). Рассмотрим теперь распределение статистики

$$z = \sum_{i=1}^{n} x_i^2,$$

где x_i — по-прежнему независимые нормальные величины с единичной дисперсией, но их средние значения отличны от нуля и

$$\mathbf{M}(x_i) = \mu_i, \qquad \sum \mu_i^2 = \lambda. \tag{24.12}$$

Запишем совместное распределение для x_i в виде

$$dF \propto \exp\left\{ -\tfrac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu})' (\boldsymbol{x} - \boldsymbol{\mu}) \right\} \prod_i dx_i.$$

С помощью ортогонального преобразования перейдем к новой системе независимых нормальных величин с единичными дисперсиями:

$$\boldsymbol{y} = \boldsymbol{B}\boldsymbol{x}.$$

Поскольку

$$\mathsf{M}(\boldsymbol{x}) = \boldsymbol{\mu}, \qquad \theta = \mathsf{M}(\boldsymbol{y}) = \boldsymbol{B}\boldsymbol{\mu},$$

то

$$\theta'\theta = \boldsymbol{\mu}'\boldsymbol{\mu} = \lambda \tag{24.13}$$

вследствие равенства $\boldsymbol{B}'\boldsymbol{B} = \boldsymbol{I}$. Сделаем теперь первые $n-1$ компонент вектора θ равными нулю. Тогда из (24.13)

$$\theta_n^2 = \lambda.$$

Таким образом, величина

$$z = \boldsymbol{x}'\boldsymbol{x} = \boldsymbol{y}'\boldsymbol{y}$$

есть сумма квадратов n независимых нормальных величин, первые $n-1$ из которых нормированы, а последняя имеет среднее $\lambda^{1/2}$ и дисперсию 1. Обозначим

$$u = \sum_{i=1}^{n-1} y_i^2, \qquad v = y_n^2,$$

где u распределена как χ^2 с $n-1$ степенями свободы. Величина y_n имеет распределение

$$dF \propto \exp\left\{-\frac{1}{2}(y_n - \lambda^{1/2})^2\right\} dy_n,$$

откуда получаем распределение для v:

$$f_1(v)\,dv \propto \frac{dv}{2v^{1/2}}\left[\exp\left\{-\frac{1}{2}(v^{1/2} - \lambda^{1/2})^2\right\} + \exp\left\{-\frac{1}{2}(-v^{1/2} - \lambda^{1/2})^2\right\}\right],$$

или

$$f_1(v)\,dv \propto v^{-1/2}\exp\left\{-\frac{1}{2}(v+\lambda)\right\}\sum_{r=0}^{\infty}\frac{(v\lambda)^r}{(2r)!}\,dv. \tag{24.14}$$

Величины v и u имеют совместное распределение

$$dG \propto f_1(v)\,f_2(u)\,dv\,du, \tag{24.15}$$

где f_2 есть χ^2-распределение с $n-1$ степенями свободы,

$$f_2(u)\,du \propto e^{-u/2}u^{(n-3)/2}\,du. \tag{24.16}$$

Подставим (24.14) и (24.16) в (24.15) и сделаем преобразование

$$z = u + v, \qquad w = \frac{u}{u+v}$$

с якобианом, равным z. Для совместного распределения z и w находим

$$dG(z, w) \propto e^{-(z+\lambda)/2}z^{(n-2)/2}w^{(n-3)/2}(1-w)^{-1/2}\sum_{r=0}^{\infty}\frac{\lambda^r z^r}{(2r)!}(1-w)^r\,dw\,dz.$$

Интегрируя теперь по w от 0 до 1, получаем маргинальное распределение величины z:

$$dH(z) \propto e^{-(z+\lambda)/2} z^{(n-2)/2} \sum_{r=0}^{\infty} \frac{\lambda^r z^r}{(2r)!} \mathrm{B}\left\{\tfrac{1}{2}(n-1), \tfrac{1}{2}+r\right\} dz. \quad (24.17)$$

Для того чтобы определить константу в (24.17), напомним, что она не зависит от λ, и положим $\lambda = 0$. Распределение (24.17) должно тогда перейти в обычное χ^2-распределение с n степенями свободы, даваемое равенством (16.1). Непостоянные множители при этом совпадают, но константа в (24.17) равна $\mathrm{B}\left\{\frac{1}{2}(n-1), \frac{1}{2}\right\} = \frac{\Gamma\{(n-1)/2\}\,\Gamma(1/2)}{\Gamma(n/2)}$, тогда как в (16.1) она равна $1/[2^{n/2}\Gamma(n/2)]$. Мы должны, следовательно, разделить (24.17) на $2^{n/2}\Gamma\{(n-1)/2\}\,\Gamma(1/2)$ и, заменяя n на ν, для любого λ окончательно имеем

$$dH(z) = \frac{e^{-(z+\lambda)/2} z^{(\nu-2)/2}}{2^{\nu/2}\Gamma\{(\nu-1)/2\}\,\Gamma(1/2)} \sum_{r=0}^{\infty} \frac{\lambda^r z^r}{(2r)!} \mathrm{B}\left\{\tfrac{1}{2}(\nu-1), \tfrac{1}{2}+r\right\} dz. \quad (24.18)$$

24.5 Распределение (24.18) называется нецентральным χ^2-распределением с ν степенями свободы и параметром нецентральности λ и обозначается иногда $\chi'^2(\nu, \lambda)$. Оно впервые введено Фишером (1928а) и изучалось Уишартом (1932), Патнайком (1949) и Тайкью (1965а). Поскольку его первые два семиинварианта равны (см. упражнение 24.1)

$$\varkappa_1 = \nu + \lambda, \qquad \varkappa_2 = 2(\nu + 2\lambda), \quad (24.19)$$

то оно следующим образом может быть приближено (центральным) χ^2-распределением. Полагая $\lambda = 0$, $\nu = \nu^*$ (в 24.19), находим первые два семиинварианта χ^2 с ν^* степенями свободы:

$$\varkappa_1 = \nu^*, \qquad \varkappa_2 = 2\nu^*. \quad (24.20)$$

Если мы приравняем между собой первые два семиинварианта величин χ'^2 и $\rho\chi^2$, где константу ρ предстоит определить, то из (24.19) и (24.20) получим

$$\nu + \lambda = \rho\nu^*, \qquad 2(\nu + 2\lambda) = 2\rho^2\nu^*,$$

так что χ'^2/ρ имеет приближенно центральное χ^2-распределение при

$$\left.\begin{aligned} \rho &= \frac{\nu + 2\lambda}{\nu + \lambda} = 1 + \frac{\lambda}{\nu + \lambda}, \\ \nu^* &= \frac{(\nu + \lambda)^2}{\nu + 2\lambda} = \nu + \frac{\lambda^2}{\nu + 2\lambda}. \end{aligned}\right\} \quad (24.21)$$

Заметим, что ν^* принимает в общем случае дробные значения.

Патнайк (1949) показал, что это приближение для ф. р. χ'^2 удовлетворительно для многих целей, но он дал и лучшие приближения, полученные из разложений в ряды Эджворта.

Если ν^* велико, то можно сделать более простое приближение, используя вместо χ^2 ее собственное приближенное распределение. Действительно, величина $(2\chi'^2/\rho)^{1/2}$ (см. **16.6**) асимптотически нормальна со средним $(2\nu^* - 1)^{1/2}$ и дисперсией 1. Кроме того, χ'^2/ρ также асимптотически нормальна с параметрами $(\nu^*, 2\nu^*)$, но сходимость в этом случае медленнее.

Если $\nu \to \infty$, то $\rho \to 1$ и $\nu^* \sim \nu$, но если $\lambda \to \infty$, то $\rho \to 2$ и $\nu^* \sim \lambda/2$.

24.6 Обобщим теперь результат из **24.4**. Предположим, что $\boldsymbol{x}$ есть n-мерный нормальный вектор со средним $\boldsymbol{\mu}$ и невырожденной матрицей рассеяния $\boldsymbol{V}$. Можно найти ортогональное преобразование $\boldsymbol{x} = \boldsymbol{By}$, которое приводит квадратичную форму $\boldsymbol{x'V^{-1}x}$ к диагональному виду $\boldsymbol{y'B'V^{-1}By} = \boldsymbol{y'Cy}$, при этом диагональные элементы матрицы $\boldsymbol{C}$ будут характеристическими числами матрицы $\boldsymbol{V}^{-1}$. Применим далее к $\boldsymbol{y'Cy}$ масштабное преобразование $\boldsymbol{y} = \boldsymbol{Dz}$, где элементы диагональной матрицы $\boldsymbol{D}$ обратны квадратным корням из соответствующих элементов $\boldsymbol{C}$, так что $\boldsymbol{D}^2 = \boldsymbol{C}^{-1}$. Таким образом, $\boldsymbol{x'V^{-1}x} = \boldsymbol{y'Cy} = \boldsymbol{z'z}$ и $\boldsymbol{z}$ есть вектор, состоящий из n независимых нормальных величин, имеющих единичную дисперсию. Среднее значение этого вектора равно $\boldsymbol{\theta}$ и удовлетворяет уравнению $\boldsymbol{\mu} = \boldsymbol{BD\theta}$. Итак, $\lambda = \boldsymbol{\theta'\theta} = \boldsymbol{\mu'V^{-1}\mu}$. Мы свели теперь задачу к случаю, рассмотренному в **24.4**. Отсюда видно, что квадратичная форма $\boldsymbol{x'V^{-1}x}$, где $\boldsymbol{x}$ — нормальный вектор с матрицей рассеяния $\boldsymbol{V}$ и средним значением $\boldsymbol{\mu}$, имеет нецентральное χ^2-распределение с n степенями свободы и параметром нецентральности $\boldsymbol{\mu'V^{-1}\mu}$. Это обобщает результат, полученный в **15.10** для многомерных нормальных величин с нулевым средним.

Грейбилл и Марсалья (1957) обобщили теоремы о распределении квадратичных форм от нормальных величин, рассматривавшиеся в **15.10—21** и в упражнениях 15.13, 15.17, на случай, когда $\boldsymbol{x}$ имеет среднее $\boldsymbol{\mu} \neq \boldsymbol{0}$. Идемпотентность матрицы является тогда необходимым и достаточным условием того, чтобы соответствующая ей квадратичная форма имела нецентральное χ^2-распределение. При такой модификации выполняются все теоремы из главы 15.

Асимптотическое распределение статистики ОП

24.7 В **18.6** было показано, что при условиях регулярности МП-оценка (которую временно обозначим t) одномерного параметра θ асимптотически достигает МГД. Согласно **17.7** ФП

имеет асимптотическое представление

$$\frac{\partial \log L}{\partial \theta} = -\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)(t-\theta), \tag{24.22}$$

или

$$L \propto \exp\left\{\frac{1}{2}\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)(t-\theta)^2\right\}. \tag{24.23}$$

Отсюда видно, что ФП сводится к нормальному распределению «асимптотически достаточной» статистики t.

В случае k-мерного вектора параметров θ матричный аналог (24.22) имеет вид

$$\frac{\partial \log L}{\partial \theta} = (t-\theta)' \boldsymbol{V}^{-1}, \tag{24.24}$$

где $\boldsymbol{V}^{-1}$ — матрица с элементами (см. **18.26**)

$$V_{ij}^{-1} = -\mathbf{M}\left(\frac{\partial^2 \log L}{\partial \theta_i \, \partial \theta_j}\right).$$

Интегрируя (24.24), получаем аналог соотношения (24.23):

$$L \propto \exp\left\{-\frac{1}{2}(t-\theta)' \boldsymbol{V}^{-1}(t-\theta)\right\}. \tag{24.25}$$

В **18.26** мы видели, что при условиях регулярности вектор МП-оценок t асимптотически имеет многомерное нормальное распределение с матрицей рассеяния $\boldsymbol{V}$. Таким образом, ФП сводится к многомерному нормальному распределению вектора t. Этот результат был строго доказан Вальдом (1943а).

Теперь можно легко установить асимптотическое распределение статистики ОП l, определенной в (24.4). В силу (24.25) задача сводится к рассмотрению отношения максимума правой части (24.25) при гипотезе H_0 к ее максимуму при H_1. Когда выполняется H_1, максимум (24.25) достигается в точке $\theta = \hat{\theta} = t$. Тогда каждая компонента вектора $(t-\theta)$ равна нулю, и мы получаем

$$L(x|\,\hat{\theta}_r, \hat{\theta}_s) \propto 1. \tag{24.26}$$

Если же имеет место гипотеза H_0, то s компонент вектора $(t-\theta)$, соответствующих θ_s, будут по-прежнему нулями, потому что максимум (24.25) достигается при $\theta_s = \hat{\hat{\theta}}_s = t_s$. При этом (24.25) имеет вид

$$L(x|\,\theta_{r0}, \hat{\hat{\theta}}_s) \propto \exp\left\{-\frac{1}{2}(t_r-\theta_{r0})' \boldsymbol{V}_r^{-1}(t_r-\theta_{r0})\right\}, \tag{24.27}$$

где индекс r указывает на то, что мы имеем дело с r-мерным распределением. Таким образом, из (24.26) и (24.27)

$$l = \frac{L(x|\,\theta_{r0}, \hat{\hat{\theta}}_s)}{L(x|\,\theta_r, \hat{\theta}_s)} = \exp\left\{-\frac{1}{2}(t_r-\theta_{r0})' \boldsymbol{V}_r^{-1}(t_r-\theta_{r0})\right\}.$$

Отсюда

$$-2\log l = (\boldsymbol{t}_r - \boldsymbol{\theta}_{r0})' \boldsymbol{V}_r^{-1} (\boldsymbol{t}_r - \boldsymbol{\theta}_{r0}).$$

Мы знаем, что $\boldsymbol{t}_r$ подчиняется многомерному нормальному распределению с матрицей рассеяния $\boldsymbol{V}_r$ и вектором средних $\boldsymbol{\theta}_r$. Итак, согласно результату из **24.6** статистика $-2\log l$ в случае проверки гипотезы, налагающей r ограничений, имеет асимптотически нецентральное χ^2-распределение с r степенями свободы и параметром нецентральности

$$\lambda = (\boldsymbol{\theta}_r - \boldsymbol{\theta}_{r0})' \boldsymbol{V}_r^{-1} (\boldsymbol{\theta}_r - \boldsymbol{\theta}_{r0}). \tag{24.28}$$

Приведенный результат принадлежит Вальду (1943a). Когда выполняется гипотеза H_0, то $\lambda = 0$ и указанное распределение сводится к центральному χ^2-распределению с r степенями свободы. Последний факт впервые был доказан Уилксом (1938a). Простое и строгое доказательство результата, относящегося к гипотезе H_0, дал К. Рой (1957). Следует подчеркнуть, что указанные результаты имеют место только тогда, когда выполнены условия, гарантирующие асимптотическую нормальность и эффективность МП-оценок.

Асимптотическая мощность критериев ОП

24.8 Результат пункта **24.7** делает возможным вычисление асимптотической функции мощности критерия ОП в тех случаях, когда выполнены условия, при которых он справедлив. Для этого мы должны сначала найти матрицу $\boldsymbol{V}_r^{-1}$, а затем интеграл

$$P = \int_{\chi'^2_\alpha(\nu,\,0)}^{\infty} d\chi'^2(\nu, \lambda), \tag{24.29}$$

где $\chi'^2_\alpha(\nu, 0)$ есть $100(1-\alpha)$-процентная точка центрального χ^2-распределения. Величина P представляет собой мощность критерия, а при $\lambda = 0$ — его размер.

Патнайк (1949) приводит таблицу значений P для $\alpha = 0{,}05$, числа степеней свободы 2(1)12(2)20 и $\lambda = 2(2)20$. При 1 степени свободы можно оценить P, используя нормальную ф. р., как показано ниже, в примере 24.3. Фикс (1949b) приводит таблицы обратной зависимости, в которых даны значения λ для числа степеней свободы 1(1)20(2)40(5)60(10)100; $\alpha = 0{,}05$, $\alpha = 0{,}01$ и P (у нее β) = 0,1(0,1)0,9 *).

Если прибегнуть к аппроксимации нецентрального распределения, приведенной в **24.5**, то, согласно (24.21) с $\nu = r$,

*) Таблицы Фикс воспроизведены в книге *Таблицы математической статистики.* (*Прим. ред.*)

соотношение (24.29) перепишется в виде

$$P = \int_{\left(\frac{r+\lambda}{r+2\lambda}\right)\chi^2_\alpha(r)}^{\infty} d\chi^2\left(r + \frac{\lambda^2}{r+2\lambda}\right), \tag{24.30}$$

где $\chi^2(r)$ — центральное χ^2-распределение с r степенями свободы, а $\chi^2_\alpha(r)$ — его $100(1-\alpha)$-процентная точка. Подстановка $\lambda = 0$ в (24.30) дает размер критерия.

Число степеней свободы в (24.30) обычно дробное, и поэтому необходима интерполяция при использовании таблиц χ^2-распределения.

Параметр нецентральности λ, определенный в (24.28), представляет собой (в предположении условий регулярности) квадратичную форму, коэффициентами которой служат элементы матрицы $\boldsymbol{V}_r^{-1}$. Из этого вытекает, что (поскольку дисперсии и ковариации имеют порядок n^{-1}) λ будет содержать множитель n, и, следовательно, мощность (24.29) будет стремиться к 1, когда n возрастает.

Пример 24.3

Проверить гипотезу H_0: $\sigma^2 = \sigma_0^2$ для нормального распределения из примера 24.1. Безусловные МП-оценки остаются теми же, так что (24.8) по-прежнему является безусловным максимумом ФП. Пусть справедлива гипотеза H_0, тогда $\hat{\hat{\mu}} = \bar{x}$ есть МП-оценка параметра μ (пример 18.2). Таким образом,

$$L\left(x \mid \hat{\hat{\mu}}, \sigma_0^2\right) = (2\pi\sigma_0^2)^{-n/2} \exp\left\{-\frac{1}{2}\frac{\sum(x-\bar{x})^2}{\sigma_0^2}\right\}. \tag{24.31}$$

Отношение (24.31) к (24.8) дает

$$l = \left(\frac{s^2}{\sigma_0^2}\right)^{n/2} \exp\left[-\frac{1}{2}n\left\{\frac{s^2}{\sigma_0^2} - 1\right\}\right],$$

поэтому

$$z = e^{-1}l^{2/n} = \frac{t}{n}e^{-t/n}, \tag{24.32}$$

где $t = ns^2/\sigma_0^2$. Величина z является монотонной функцией от l, но она не монотонна относительно t/n. Ее производная равна

$$\frac{dz}{dt} = \frac{1}{n}\left(1 - \frac{t}{n}\right)e^{-t/n}.$$

Следовательно, z возрастает при $t < n$ до максимума в точке $t = n$ и затем убывает. Неравенство $l \leqslant c_\alpha$ эквивалентно тому,

что $t \leqslant a_\alpha$, $t \geqslant b_\alpha$, где a_α, b_α, согласно (24.32), определяются из соотношений

$$\left.\begin{aligned} \mathbf{P}\{t \leqslant a_\alpha\} + \mathbf{P}\{t \geqslant b_\alpha\} &= \alpha, \\ a_\alpha e^{-a_\alpha/n} &= b_\alpha e^{-b_\alpha/n}. \end{aligned}\right\} \tag{24.33}$$

Поскольку при гипотезе H_0 статистика t имеет χ^2-распределение с $n-1$ степенями свободы, то с помощью таблиц этого распределения можно решить систему уравнений (24.33).

Рассмотрим теперь приближенное распределение величины

$$-2\log l = (t-n) - n\log\left(\frac{t}{n}\right).$$

Так как $\mathbf{M}t = n-1$, $\mathbf{D}t = 2(n-1)$, то

$$-2\log l = (t-n) - n\log\left(1 + \frac{t-n}{n}\right) =$$

$$= (t-n) - n\sum_{r=1}^{\infty}(-1)^{r-1}\left(\frac{t-n}{n}\right)^r / r =$$

$$= (t-n) - n\left\{\frac{t-n}{n} - \frac{(t-n)^2}{2n^2} + o(n^{-2})\right\} = \frac{(t-n)^2}{2n} + o(n^{-1}). \tag{24.34}$$

Мы видели в **16.6**, что χ^2-величина с $n-1$ степенями свободы при $n \to \infty$ асимптотически нормальна со средним $(n-1)$ и дисперсией $2(n-1)$ или, что эквивалентно, величина $(t-n)/(2n)^{1/2}$ стремится к стандартной нормальной величине. Следовательно, первое слагаемое в (24.34), представляющее собой ее квадрат, имеет χ^2-распределение с одной степенью свободы. Именно это распределение величины $-2\log l$ дает общий результат из **24.7** в случае, когда выполняется H_0. Этот результат говорит также, что, когда гипотеза H_0 неверна, $-2\log l$ подчиняется нецентральному χ^2-распределению с одной степенью свободы и параметром нецентральности, равным согласно (24.28)

$$\lambda = -\mathbf{M}\left\{\frac{\partial^2 \log L}{\partial(\sigma^2)^2}\right\}(\sigma^2 - \sigma_0^2)^2 = \frac{n}{2\sigma^4}(\sigma^2 - \sigma_0^2)^2 = \frac{n}{2}\left(1 - \frac{\sigma_0^2}{\sigma^2}\right)^2.$$

Таким образом, в данном случае, когда $r = 1$, выражение (24.30) для приближенной мощности критерия ОП имеет вид

$$P = \int_{\left(\frac{1+\lambda}{1+2\lambda}\right)\chi^2_\alpha(1)}^{\infty} d\chi^2\left(1 + \frac{\lambda^2}{1+2\lambda}\right). \tag{24.35}$$

Для иллюстрации оценим P для одного значения λ и удобным

образом выбранного n. Задавая $\chi^2_\alpha(1) = 3{,}84$, получаем критерий уровня 0,05. Пусть $\sigma^2 = \sigma_1^2 = 1{,}25\sigma_0^2$ — альтернативная гипотеза. Тогда имеем $\lambda = 0{,}02\,n$ и, полагая $n = 50$, находим $\lambda = 1$. Из соотношения (24.35) вытекает, что

$$P = \int_{2{,}56}^{\infty} d\chi^2\left(\frac{4}{3}\right).$$

Используя *Biometrika Tables*, с помощью простой интерполяции между 1 и 2 степенями свободы определяем, что приближенно $P = 0{,}166$. Точное значение мощности может быть получено из нормальной ф. р. Это есть мощность критерия размера α с равными «хвостами» против альтернативной гипотезы, состоящей в смещении среднего на $\lambda^{1/2} = 1$ среднеквадратичное отклонение от среднего при H_0, т. е. доля альтернативного распределения, лежащая вне интервала $(-2{,}96, +0{,}96)$ среднеквадратичных отклонений от его среднего. Таблицы нормального распределения дают значение $P = 0{,}170$. Аппроксимация для функции мощности получилась, таким образом, вполне хорошей.

Более точные аппроксимации для распределения статистики ОП

24.9 Ограничиваясь теперь распределением статистики l при гипотезе H_0, можно искать более точные приближения, чем асимптотический результат из **24.7**. В **24.3** было указано, что если мы хотим найти χ^2-приближение для распределения функции от l, то полезно рассмотреть величину $w = -2c \log l$ и подобрать константу c так, чтобы улучшить аппроксимацию.

Для этого проще всего было бы найти математическое ожидание w и подобрать такое c, при котором оно равно математическому ожиданию величины χ^2 с r степенями свободы, т. е. $\mathbf{M}(w) = r$. Аппроксимация такого сорта впервые была получена Бартлеттом (1937). Общий способ определения константы c предложил Лоли (1956), который для исследования моментов величины $-2\log l$ существенно использует методы из **20.15**. Если

$$\mathbf{M}(-2\log l) = r\left\{1 + \frac{a}{n} + O\left(\frac{1}{n^2}\right)\right\}, \tag{24.36}$$

то, как показал Лоли, полагая

$$w_1 = -2\left(\frac{1}{1 + a/n}\right)\log l \quad \text{либо} \quad w_2 = -2\left(1 - \frac{a}{n}\right)\log l, \tag{24.37}$$

мы не только получаем, что

$$\mathbf{M}(w) = r + O\left(\frac{1}{n^2}\right),$$

как это сразу следует из (24.36) и (24.37), но и что все семиинварианты статистики w совпадают, с точностью до n^{-1}, с семиинвариантами величины χ^2, имеющей r степеней свободы. Простая масштабная поправка, которая согласует среднее w с правильным значением, является, таким образом, несомненным улучшением.

Если требуются более точные приближения, то их можно получить с помощью метода, принадлежащего Боксу (1949). Бокс приводит улучшенные χ^2-аппроксимации (см. **42.11**, т. 3) и показывает, как найти функцию от $-2\log l$, имеющую распределение дисперсионного отношения, а также выводит асимптотическое разложение для ее распределения в терминах неполных гамма-функций.

Пример 24.4

Пусть имеется k независимых выборок объемов n_i ($i = 1, 2, \ldots, k$; $n_i \geqslant 2$), извлеченных из нормальных совокупностей со средними μ_i и дисперсиями σ_i^2. Нужно проверить гипотезу

$$H_0\colon\ \sigma_1^2 = \sigma_2^2 = \ \ldots\ = \sigma_k^2;$$

это сложная гипотеза, налагающая $r = k - 1$ ограничений

$$\frac{\sigma_2^2}{\sigma_1^2} = \frac{\sigma_3^2}{\sigma_1^2} = \ \ldots\ = \frac{\sigma_k^2}{\sigma_1^2} = 1$$

и имеющая $s = k + 1$ степеней свободы. Обозначим σ^2 общее неизвестное значение дисперсии.

Безусловный максимум ФП получается, как и в примере 24.1, если положить

$$\hat{\mu}_i = \bar{x}_i, \qquad \hat{\sigma}_i^2 = \frac{1}{n_i}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2 = s_i^2.$$

Он равен

$$L(x|\,\hat{\mu}_1, \ldots, \hat{\mu}_k, \hat{\sigma}_1^2, \ldots, \hat{\sigma}_k^2) = (2\pi)^{-n/2}\prod_{i=1}^{k}(s_i^2)^{-n_i/2}e^{-n/2}, \quad (24.38)$$

где

$$n = \sum_{i=1}^{k} n_i.$$

При гипотезе H_0 МП-оценки средних и общей дисперсии имеют вид

$$\hat{\hat{\mu}}_i = \bar{x}_i, \qquad \hat{\hat{\sigma}}_2^2 = \frac{1}{n}\sum_{i=1}^{k} n_i s_i^2 = s^2,$$

откуда

$$L(x|\,\hat{\hat{\mu}}_1, \ldots, \hat{\hat{\mu}}_k, \hat{\hat{\sigma}}^2) = (2\pi)^{-n/2}(s^2)^{-n/2}e^{-n/2}. \quad (24.39)$$

Из (24.4), (24.38) и (24.39) имеем

$$l = \prod_{i=1}^{k} \left(\frac{s_i^2}{s^2} \right)^{n_i/2}, \tag{24.40}$$

так что

$$-2 \log l = n \log (s^2) - \sum_{i=1}^{k} n_i \log (s_i^2). \tag{24.41}$$

Когда выполняется гипотеза H_0, каждая из статистик $\left(\frac{n_i s_i^2}{2\sigma^2} \right)$ имеет гамма-распределение с параметром $\frac{1}{2}(n_i - 1)$, а их сумма $\frac{ns^2}{2\sigma^2}$ — такое же распределение уже с параметром $\sum_{i=1}^{k} \frac{1}{2}(n_i - 1) = \frac{1}{2}(n-k)$. Для величины x, имеющей гамма-распределение с параметром p, имеем

$$\mathbf{M}\{\log (ax)\} = \frac{1}{\Gamma(p)} \int_0^\infty \log (ax)\, e^{-x} x^{p-1}\, dx = \log a + \frac{d}{dp} \log \Gamma(p),$$

откуда, используя ряд Стирлинга (3.63), получаем

$$\mathbf{M}\{\log (ax)\} = \log a + \log p - \frac{1}{2p} - \frac{1}{12p^2} + O\left(\frac{1}{p^3} \right). \tag{24.42}$$

Пользуясь (24.42), получаем из (24.41)

$$\begin{aligned}
&\mathbf{M}\{-2 \log l\} = \\
&= n \left\{ \log \left(\frac{2\sigma^2}{n} \right) + \log \left\{ \frac{1}{2}(n-k) \right\} - \frac{1}{(n-k)} - \frac{1}{3(n-k)^2} + O\left(\frac{1}{n^3} \right) \right\} - \\
&\quad - \sum_{i=1}^{k} n_i \left\{ \log \left(\frac{2\sigma^2}{n_i} \right) + \log \left\{ \frac{1}{2}(n_i - 1) \right\} - \right. \\
&\quad \left. - \frac{1}{n_i - 1} - \frac{1}{3(n_i-1)^2} + O\left(\frac{1}{n_i^3} \right) \right\} = \\
&= n \left\{ \log \left(1 - \frac{k}{n} \right) - \frac{1}{(n-k)} - \frac{1}{3(n-k)^2} + O\left(\frac{1}{n^3} \right) \right\} - \\
&\quad - \sum_{i=1}^{k} n_i \left\{ \log \left(1 - \frac{1}{n_i} \right) - \frac{1}{(n_i - 1)} - \frac{1}{3(n_i-1)^2} + O\left(\frac{1}{n_i^3} \right) \right\} = \\
&= (k-1) + \left[\left(\sum_{i=1}^{k} \frac{1}{n_i - 1} - \frac{k}{n-k} \right) + \frac{1}{2} \left(\sum_{i=1}^{k} \frac{1}{n_i} - \frac{k^2}{n} \right) + \right. \\
&\quad \left. + \frac{1}{3} \left\{ \sum_{i=1}^{k} \frac{n_i}{(n_i-1)^2} - \frac{n}{(n-k)^2} \right\} \right] + O\left(\frac{1}{N^3} \right),
\end{aligned} \tag{24.43}$$

где N обозначает любую из величин n_i или n. Мы можем теперь улучшить χ^2-приближение, полагая, в соответствии с (24.36), выражение в квадратных скобках (24.43) равным $(k-1)\frac{a}{n}$.

Рассмотрим теперь модификацию Бартлетта (1937) для статистики ОП (24.40), в которой n_i повсюду заменяется на «число степеней свободы» $n_i - 1 = \nu_i$, так что n заменяется на $\nu = \sum_{i=1}^{k} (n_i - 1) = n - k$. Мы обозначим ее

$$l^* = \prod_{i=1}^{k} \left(\frac{s_i^2}{s^2}\right)^{\frac{1}{2}\nu_i},$$

где теперь

$$s_i^2 = \frac{1}{\nu_i} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2, \quad s^2 = \frac{1}{\nu} \sum_{i=1}^{k} \nu_i s_i^2.$$

Таким образом,

$$-2 \log l^* = \nu \log s^2 - \sum_{i=1}^{k} \nu_i \log s_i^2. \tag{24.44}$$

В примере 24.6 мы увидим, что статистика l^* обладает тем преимуществом перед l, что соответствующий ей критерий несмещен при любых значениях n_i. Если мы повторим переход от (24.42) к (24.43), то найдем, что

$$\mathbf{M}(-2 \log l^*) =$$

$$= -\nu \left\{\frac{1}{\nu} + \frac{1}{3\nu^2} + O\left(\frac{1}{\nu^3}\right)\right\} + \sum_{i=1}^{k} \nu_i \left\{\frac{1}{\nu_i} + \frac{1}{3\nu_i^2} + O\left(\frac{1}{\nu_i^3}\right)\right\} =$$

$$= (k-1) + \frac{1}{3}\left(\sum_{i=1}^{k} \frac{1}{\nu_i} - \frac{1}{\nu}\right) + O\left(\frac{1}{\nu_i^3}\right). \tag{24.45}$$

Из (24.37) и (24.45) следует, что величину $-2 \log l^*$, определенную в (24.44), нужно разделить на

$$1 + \frac{1}{3(k-1)}\left(\sum_{i=1}^{k} \frac{1}{\nu_i} - \frac{1}{\nu}\right),$$

чтобы получить более точное приближение к χ^2-распределению с $k-1$ степенями свободы.

Критерии ОП в случае, когда область зависит от параметра

24.10 Асимптотическое распределение статистики ОП, данное в **24.7**, существенно опирается на условия регулярности, необходимые для того, чтобы установить асимптотическую нормальность МП-оценок. Эти условия нарушаются (как мы видели в примере 18.5), если область определения исходного распределения зависит от параметра θ. Что же в таком случае можно сказать о распределении ОП-статистики? Замечательный факт, доказанный Хоггом (1956), состоит в том, что для определенных гипотез, касающихся указанных распределений, статистика $-2\log l$ имеет *в точности* χ^2-распределение, но с $2r$ степенями свободы, т. е. число степеней свободы вдвое больше количества ограничений, налагаемых гипотезой.

24.11 Выведем сначала некоторые предварительные результаты, касающиеся прямоугольных распределений. Пусть k независимых величин z_i распределены по закону

$$dF = dz_i, \quad 0 \leqslant z_i \leqslant 1. \tag{24.46}$$

Величина

$$t_i = -2\log z_i$$

имеет, как легко видеть, распределение

$$dG = \frac{1}{2}\exp\left(-\frac{1}{2}t_i\right)dt_i, \quad 0 \leqslant t_i \leqslant \infty,$$

являющееся χ^2-распределением с двумя степенями свободы. Поэтому сумма k таких величин

$$t = \sum_{i=1}^{k} t_i = -2\sum_{i=1}^{k}\log z_i = -2\log\prod_{i=1}^{k} z_i$$

подчиняется χ^2-распределению с $2k$ степенями свободы.

Из (14.1) следует, что наибольшее наблюдение $y_{(n_i)}$ среди n_i независимых наблюдений из прямоугольного распределения на (0, 1) имеет распределение

$$dH = n_i y_{(n_i)}^{n_i-1}\,dy_{(n_i)}, \quad 0 \leqslant y_{(n_i)} \leqslant 1, \tag{24.47}$$

и, следовательно, величина $y_{(n_i)}^{n_i} = z_i$ распределена равномерно по закону (24.46). Таким образом, в случае k независимых выборок объемов n_i статистика $t = -2\log\prod_{i=1}^{k} y_{(n_i)}^{n_i}$ имеет χ^2-распределение с $2k$ степенями свободы.

Рассмотрим теперь распределение наибольшего среди k наибольших значений $y_{(n_i)}$. Поскольку все наблюдения независимы,

то это попросту есть наибольшее среди $n = \sum_{i=1}^{k} n_i$ наблюдений, взятых из исходного прямоугольного распределения. Если обозначим указанное наибольшее значение через $y_{(n)}$, то величина $-2\log y_{(n)}^n$, согласно сказанному выше, будет иметь χ^2-распределение с двумя степенями свободы. Докажем теперь независимость статистик $y_{(n)}$ и $u = \prod_{i=1}^{k} y_{(n_i)}^{n_i} \Big/ y_{(n)}^n$. Пусть исходное прямоугольное распределение сосредоточено на интервале $(0, \theta)$. Тогда согласно (24.47) совместная плотность величин $y_{(n_i)}$ равна

$$f = \prod_{i=1}^{k} \left\{ n_i y_{(n_i)}^{n_i - 1} / \theta^{n_i} \right\} = \frac{1}{\theta^n} \prod_{i=1}^{k} n_i y_{(n_i)}^{n_i}.$$

Как показано в **17.40**, $y_{(n)}$ — достаточная статистика для θ, а в силу **23.12** ее распределение полно. Таким образом, согласно упражнению 23.7, для того чтобы установить независимость величины u от полной достаточной статистики $y_{(n)}$ нам нужно только заметить, что распределение u не зависит от параметра θ.

Величины $y_{(n)}$ и u, а следовательно, и $y_{(n)}^n$ и u независимы. Обозначим $\varphi_1(t)$ х. ф. величины $-2\log y_{(n)}^n$, $\varphi_2(t)$ — х. ф. $-2\log \sum_{i=1}^{k} y_{(n_i)}^{n_i}$ и $\varphi(t)$ — х. ф. $-2\log u$. Из предыдущего имеем

$$(-2\log u) + (-2\log y_{(n)}^n) = -2\log \prod_{i=1}^{k} y_{(n_i)}^{n_i}.$$

Используя предыдущие результаты о χ^2-распределениях и тот факт, что х. ф. суммы независимых величин равна произведению их х. ф. (см. **7.18**) получаем

$$\varphi(t)(1-2it)^{-1} = (1-2it)^{-k},$$

откуда

$$\varphi(t) = (1-2it)^{-(k-1)}.$$

Следовательно, статистика $-2\log u$ имеет χ^2-распределение с $2(k-1)$ степенями свободы.

Итак, мы установили, что если имеется k независимых величин $y_{(n_i)}$, распределенных по закону (24.47), то величина $-2\log \prod_{i=1}^{k} y_{(n_i)}^{n_i}$ имеет χ^2-распределение с $2k$ степенями свободы. Далее, если $y_{(n)}$ — наибольшее среди $y_{(n_i)}$, то статистика $-2\log \left\{ \sum_{i=1}^{k} y_{(n_i)}^{n_i} \Big/ y_{(n)}^n \right\}$ подчиняется χ^2-распределению с $2(k-1)$ степенями свободы.

24.12 Рассмотрим теперь два типа ситуаций, в которых существует одномерная достаточная статистика для θ, когда область определения зависит от параметра θ. Возьмем случай, где только один конец (например, верхний) зависит от θ. Согласно результатам **17.40** плотность имеет вид

$$f(x \mid \theta) = g(x)/h(\theta), \qquad a \leqslant x \leqslant \theta. \tag{24.48}$$

Предположим теперь, что имеется $k(\geqslant 1)$ генеральных совокупностей с плотностями $f(x_i \mid \theta_i)$, и мы хотим, основываясь на выборках объемов n_i $(i = 1, 2, \ldots, k)$, проверить простую гипотезу

$$H_0\colon\ \theta_1 = \theta_2 = \ldots = \theta_k = \theta_0,$$

налагающую k ограничений. Построим критерий ОП для H_0. Наибольшее наблюдение $x_{(n_i)}$ (см. упражнение 18.1) является безусловной МП-оценкой параметра θ_i. Таким образом,

$$L(x \mid \hat{\theta}_1, \ldots, \hat{\theta}_k) = \prod_{i=1}^{k} \prod_{j=1}^{n_i} \{g(x_{ij})/[h(x_{(n_i)})]^{n_i}\}. \tag{24.49}$$

Поскольку H_0 — простая гипотеза, то, когда она выполняется, значение ФП определено однозначно и не нужно искать условную МП-оценку. Имеем

$$L(x \mid \theta_0, \theta_0, \ldots, \theta_0) = \prod_{i=1}^{k} \prod_{j=1}^{n_i} \{g(x_{ij})\}/[h(\theta_0)]^n.$$

Следовательно, статистика ОП равна

$$l = \frac{L(x \mid \theta_0, \ldots, \theta_0)}{L(x \mid \hat{\theta}_1, \ldots, \hat{\theta}_k)} = \prod_{i=1}^{k} \left[\frac{h(x_{(n_i)})}{h(\theta_0)}\right]^{n_i}. \tag{24.50}$$

Если H_0 выполнена, то величина $y_i = h(x_{(n_i)})/h(\theta_0)$ представляет собой вероятность того, что наблюдение меньше или равно $x_{(n_i)}$. Она сама является случайной величиной, распределение которой находится по распределению $x_{(n_i)}$:

$$dF = n_i y_i^{n_i - 1}\, dy_i, \qquad 0 \leqslant y_i \leqslant 1,$$

и имеет вид (24.47). Таким образом, согласно результату предыдущего пункта статистика

$$-2 \log \prod_{i=1}^{k} y_i^{n_i} = -2 \log l$$

подчиняется χ^2-распределению с $2k$ степенями свободы.

24.13 Рассмотрим теперь для $k \geqslant 2$ генеральных совокупностей сложную гипотезу

$$H_0\colon\ \theta_1 = \theta_2 = \ldots = \theta_k,$$

которая налагает $k-1$ ограничений, поскольку общее значение параметра θ не фиксировано. Как и прежде, безусловный максимум ФП дается соотношением (24.49). Максимум при гипотезе H_0 равен $L(x \mid \hat{\hat{\theta}}, \hat{\hat{\theta}}, \ldots, \hat{\hat{\theta}})$, где $\hat{\hat{\theta}}$ — МП-оценка для θ по объединенной выборке, равная $x_{(n)}$. Таким образом, находим статистику ОП

$$l = \frac{L(x \mid \hat{\hat{\theta}}, \ldots, \hat{\hat{\theta}})}{L(x \mid \hat{\theta}_1, \ldots, \hat{\theta}_k)} = \prod_{i=1}^{k} \frac{[h(x_{(n_i)})]^{n_i}}{[h(x_n)]^n}. \tag{24.51}$$

Перепишем ее в виде

$$l = \prod_{i=1}^{k} \left[\frac{h(x_{(n_i)})}{h(\theta)}\right]^{n_i} \Big/ \left[\frac{h(x_{(n)})}{h(\theta)}\right]^{n},$$

где θ — общее неизвестное значение параметра θ_i. Отсюда видно, что в обозначениях предыдущего пункта

$$l = \left[\prod_{i=1}^{k} y_{(n_i)}^{n_i}\right] \Big/ y_{(n)}^{n},$$

и согласно **24.11** статистика $-2\log l$ в данном случае подчиняется χ^2-распределению с $2(k-1)$ степенями свободы.

24.14 Если оба конца интервала зависят от параметра θ и для θ существует одномерная достаточная статистика, то согласно **17.41**

$$f(x \mid \theta) = g(x)/h(\theta), \qquad \theta \leqslant x \leqslant b(\theta), \tag{24.52}$$

где $b(\theta)$ должна быть монотонно убывающей функцией от θ. Пусть имеется k таких совокупностей, $f(x_i \mid \theta_i)$, и на основе выборок объемов n_i снова проверяется простая гипотеза

$$H_0\colon\ \theta_1 = \theta_2 = \ldots = \theta_k = \theta_0.$$

Безусловной МП-оценкой параметра θ_i является достаточная статистика

$$t_i = \min\{x_{(1i)},\ b^{-1}(x_{(ni)})\},$$

где $x_{(1i)}$ и $x_{(ni)}$ — соответственно наименьшее и наибольшее наблюдения в i-й выборке. При гипотезе H_0 ФП равна $L(x \mid \theta_0, \ldots, \ldots, \theta_0)$. Отсюда получаем статистику ОП

$$l = \frac{L(x \mid \theta_0, \ldots, \theta_0)}{L(x \mid \hat{\theta}_1, \ldots, \hat{\theta}_k)} = \prod_{i=1}^{k} \left[\frac{h(t_i)}{h(\theta_0)}\right]^{n_i}. \tag{24.53}$$

Точно так же, как в случае соотношения (24.50), убеждаемся, что

$$l = \prod_{i=1}^{k} y_i^{n_i},$$

где y_i подчиняются распределению (24.47), и, следовательно, величина $-2 \log l$ снова имеет χ^2-распределение с $2k$ степенями свободы.

Аналогичным образом, в случае сложной гипотезы с $k-1$ ограничениями ($k \geqslant 2$)

$$H_0\colon\ \theta_1 = \theta_2 = \ldots = \theta_k$$

так же, как в **24.13**, находим, что статистика ОП равна

$$l = \frac{L(x \mid \hat{\hat{\theta}}, \ldots, \hat{\hat{\theta}})}{L(x \mid \hat{\theta}_1, \ldots, \hat{\theta}_k)} = \prod_{i=1}^{k} [h(t_i)]^{n_i} / [h(t)]^n,$$

где $t = \min\{t_i\}$ — МП-оценка $\hat{\hat{\theta}}$ по объединенной выборке. Записывая

$$l = \prod_{i=1}^{k} \left[\frac{h(t_i)}{h(\theta)}\right]^{n_i} \Big/ \left[\frac{h(t)}{h(\theta)}\right]^n,$$

мы вновь приводим l к требуемому в **24.11** виду. Отсюда вытекает, что статистика $-2 \log l$ имеет χ^2-распределение с $2(k-1)$ степенями свободы.

24.15 Таким образом, мы получили точные χ^2-распределения для двух классов проверяемых гипотез, касающихся распределений, у которых концы интервалов зависят от параметра. Упражнения 24.8 и 24.9 содержат дальнейшие примеры критериев ОП, одного точного и одного асимптотического, в которых статистика $-2 \log l$ имеет χ^2-распределение с удвоенным, по сравнению с количеством ограничений в проверяемой гипотезе, числом степеней свободы. Заметим, что указанные χ^2-распределения возникают не из-за сходимости МП-оценок к многомерной нормальности, как это было с асимптотическими результатами для «регулярных» ситуаций в **24.7**, а вследствие выявленной в **24.11** тесной связи между прямоугольным и χ^2-распределениями. Одно из следствий этого различия состоит в том, что функции мощности рассматриваемых критериев совершенно отличаются от полученных в **24.8** с помощью нецентрального χ^2-распределения.

Для простой гипотезы Барр (1966) нашел функцию мощности критерия, основанного на статистике ОП (24.50), и доказал его несмещенность. Он показал, что при $k = 1$ этот критерий является РНМ (см. пример 22.6 и условие (22.37)); если же

$k > 1$, то РНМ критерия не существует, а критерий ОП не является даже РНМН. В случае сложной гипотезы, приведенной в 24.13, при $k = 2$ с помощью функции мощности критерия ОП со статистикой (24.51) доказывается, что этот критерий ОП является РНМН.

Свойства критериев ОП

24.16 До сих пор мы всецело занимались нахождением распределения статистики ОП или некоторой функции от нее. Теперь нас будут интересовать свойства критериев ОП, в частности их несмещенность и свойства оптимальности. Однако вначале рассмотрим более слабое свойство, а именно свойство состоятельности, которое будет сейчас определено.

Состоятельность критерия

24.17 Критерий для проверки гипотезы H_0 против класса альтернатив H_1 называется состоятельным, если при любой альтернативе из H_1 вероятность отвергнуть H_0 стремится к 1, когда объем выборки (выборок) стремится к бесконечности. Пусть w есть критическая область, а x — выборочная точка. Тогда определение можно записать так:

$$\lim_{n\to\infty} \mathsf{P}\{x \in w \mid H_1\} = 1. \tag{24.54}$$

Простая и естественная идея состоятельности критерия была впервые введена Вальдом и Волфовицем (1940). Когда возрастает число наблюдений, кажется совершенно разумным требовать, чтобы любой критерий, заслуживающий рассмотрения, отклонял ложную гипотезу с возрастающей, а в пределе с полной достоверностью. Состоятельность критерия является столь же естественным свойством, как и состоятельность оценки (17.7), и представляет собой в некотором смысле обобщение последней. Действительно, если критерий, касающийся параметра θ, основан на статистике, которая является состоятельной оценкой для θ, то совершенно очевидно, что критерий также будет состоятельным. Однако состоятельный критерий можно получить и с помощью несостоятельной оценки. Например, если t стремится по вероятности к $a\theta$, то t дает состоятельный критерий для гипотезы о параметре θ. Вообще, для состоятельности критерия достаточно, чтобы статистика критерия, рассматриваемая как оценка, стремилась по вероятности к некоторой взаимно однозначной функции от θ.

Поскольку условием несмещенности критерия размера α (см. (23.53)) является неравенство

$$\mathsf{P}\{x \in w \mid H_1\} \geqslant \alpha, \tag{24.55}$$

то из (24.54) и (24.55) понятно, что состоятельный критерий будет терять смещение (если оно было) при $n \to \infty$. Однако несмещенный критерий не обязан быть состоятельным *).

Состоятельность и несмещенность критериев ОП

24.18 В **18.10** и **18.22** мы видели, что при некоторых очень слабых ограничениях МП-оценка $\hat{\theta}$ вектора θ состоятельна, хотя в общем случае она может не быть таковой Если мы имеем дело с ситуацией, в которой все МП-оценки состоятельны, то из определения статистики ОП (24.4) видно, что с ростом объема выборки

$$l \to \frac{L(x \mid \theta_{r0}, \theta_s)}{L(x \mid \theta_r, \theta_s)}, \tag{24.56}$$

где θ_r, θ_s — истинные значения этих параметров, а θ_{r0} — задаваемое гипотезой значение параметра θ_r. Таким образом, когда выполняется H_0, $l \to 1$ по вероятности и, следовательно, граница c_α критической области (24.6) будет приближаться к 1. При невыполнении H_0 предельным значением l в (24.56) будет некоторая константа k, удовлетворяющая (см. (18.20)) неравенству $0 \leqslant k < 1$, и тогда имеем

$$\mathsf{P}\{l \leqslant c_\alpha\} \to 1, \tag{24.57}$$

т. е. критерий ОП состоятелен.

В **24.8**, исходя из приближенного вида функции мощности, мы получили, что при условиях регулярности критерии ОП состоятельны. В **24.15** была выведена состоятельность для другого случая, не содержащегося в **24.8**. Оба этих примера представляют собой частные случаи полученного сейчас результата.

24.19 Обращаясь к вопросу о несмещенности, вспомним предпоследнее предложение из **24.17**, которое в совокупности с результатом из **24.18** обеспечивает асимптотическую несмещенность большинства критериев ОП. Само по себе это не очень утешительно (хотя это было бы так, если бы при разумных ограничениях можно было показать, что максимальное смещение всегда мало), потому что интуитивно кажется разумным требовать, чтобы условие несмещенности критерия выполнялось для выборок всех объемов. Пример 24.5 показывает, что один важный критерий ОП является смещенным.

Пример 24.5

Рассмотрим снова гипотезу H_0 из примера 24.3. Критической областью критерия ОП служат «хвосты» χ^2_{n-1}-распределения величины $t = ns^2/\sigma_0^2$, определяемые уравнениями (24.33).

*) См. замечание в **17.9** о состоятельных и асимптотически несмещенных оценках.

В примерах 23.12 и 23.14 мы видели, что несмещенный (в действительности РНМН) критерий для H_0 определяется из соотношений

$$\left.\begin{aligned} &\mathsf{P}\{t \leqslant a_\alpha\} + \mathsf{P}\{t \geqslant b_\alpha\} = \alpha, \\ &a_\alpha^{(n-1)/2} \exp(-a_\alpha/2) = b_\alpha^{(n-1)/2} \exp(-b_\alpha/2). \end{aligned}\right\} \qquad (24.58)$$

Из сравнения (24.58) и (24.33) вытекает, что они дают одинаковый результат только при условии $a_\alpha = b_\alpha$, которое выполняется лишь в тривиальном случае $a_\alpha - b_\alpha = 0$, $\alpha = 1$. Во всех других случаях эти критерии имеют разные критические области, причем у критерия ОП значения a_α и b_α больше, чем у несмещенного критерия, т. е. у него бóльшая доля величины α сконцентрирована на нижнем «хвосте». Рассмотрим альтернативные значения σ^2 бóльшие, чем σ_0^2, для которых распределение статистики t немного смещается в сторону бóльших значений t по сравнению с распределением при гипотезе H_0. Нетрудно проверить, что при таких σ^2 уменьшение вероятности, соответствующей критической области критерия ОП, вызванное увеличением b_α, будет превосходить ее прирост вследствие бóльшего значения a_α. Таким образом, критерий ОП является смещенным.

На примере 23.12 можно увидеть, что рассмотренный там критерий с равными «хвостами» имеет значения a_α, b_α, слишком малые для несмещенности, тогда как значения a_α, b_α, соответствующие критерию ОП, для этого слишком велики. Таким образом, две более или менее интуитивно приемлемые критические области как бы «ограничивают» с двух сторон критическую область несмещенного критерия.

Если в (24.33) заменить n на $n-1$, то получим в точности соотношения (24.58) , соответствующие несмещенному критерию, что подтверждает общее утверждение о том, что критерий ОП асимптотически теряет свое смещение. Это наводит на мысль найти причину смещения. Если при построении статистики ОП в примере 24.3 мы подправим безусловную МП-оценку параметра σ^2 так, чтобы она была несмещенной, то s^2 заменится на $\left(\frac{n}{n-1}\right) s^2$ и подправленный критерий ОП будет несмещенным. Итак, смещенность МП-оценки лежит в основе смещенности критерия ОП.

Несмещенные критерии для параметров сдвига и масштаба

24.20 Пример 24.5 подсказывает, что при построении критерия ОП полезно подправить используемые МП-оценки с тем, чтобы они были несмещенными. Подтверждением сказанному служит пример 24.4, где было установлено, что подправленная

статистика ОП l^* порождает несмещенный критерий. Сейчас мы докажем это утверждение, используя метод, принадлежащий Питмэну (1939b).

Если проверяемая гипотеза касается k параметров сдвига θ_i ($i = 1, 2, \ldots, k$), то совместное распределение случайных величин запишется в виде

$$dF = f(x_1 - \theta_1,\ x_2 - \theta_2,\ \ldots,\ x_k - \theta_k)\,dx_1 \ldots dx_k. \quad (24.59)$$

Мы хотим проверить гипотезу

$$H_0:\ \theta_1 = \theta_2 = \ldots = \theta_k. \quad (24.60)$$

Любая интуитивно подходящая для построения критерия статистика t должна удовлетворять условию инвариантности

$$t(x_1,\ x_2,\ \ldots,\ x_k) = t(x_1 - \lambda,\ x_2 - \lambda,\ \ldots,\ x_k - \lambda). \quad (24.61)$$

Следовательно, без потери общности можно положить общее значение θ_i в (24.60) равным нулю. Предположим, что $t > 0$ и что критическая область w_0 размера α, основанная на распределении статистики t, определяется неравенством

$$t \leqslant c_\alpha. \quad (24.62)$$

Если бы какие-нибудь из этих предположений не выполнялись, то мы могли бы перейти к некоторой функции от t, для которой они имели бы место.

Вследствие свойства инвариантности (24.61) статистика t должна быть постоянной (в k-мерном пространстве W) на любой прямой L, параллельной вектору V, определяемому уравнениями $x_1 = x_2 = \ldots = x_k$. Таким образом, область w_0 будет лежать вне гиперцилиндра с осью, параллельной V. При гипотезе H_0 вероятность области w_0 равна размеру критерия

$$\alpha = \int_{w_0} dF(x_1,\ x_2,\ \ldots,\ x_k). \quad (24.63)$$

Если же H_0 неверна, то вероятность области w_0 равна мощности критерия

$$1 - \beta = \int_{w} dF(x_1 - \theta_1,\ x_2 - \theta_2,\ \ldots,\ x_k - \theta_k) = \int_{w_1} dF(x_1,\ x_2,\ \ldots,\ x_k), \quad (24.64)$$

где область w_1 получена из w_0 параллельным переносом в пространстве W. Определим интеграл по любой прямой L, параллельной V,

$$P(L) = \int_L f(x_1,\ x_2,\ \ldots,\ x_k)\,d\bar{x}. \quad (24.65)$$

Поскольку $\bar{x}$ пробегает произвольную прямую L, а совокупность всех прямых L совпадает с пространством W, то

$$\int P(L)\,dL = \int \left\{ \int_L f\,d\bar{x} \right\} dL = \int \ldots \int f\,dx_1 \ldots dx_k = 1. \qquad (24.66)$$

Определим теперь область w_0 как объединение всех прямых L, для которых статистика $P(L)$ не превосходит некоторой постоянной h. Тогда $P(L)$ будет больше h на любой прямой L, принадлежащей w_1 и не принадлежащей w_0. Следовательно, из (24.63), (24.64) и (24.66) вытекает

$$\alpha = \int_{w_0} dF \leqslant \int_{w_1} dF = 1 - \beta,$$

поэтому данный критерий является несмещенным. Нам осталось определить статистику критерия t так, чтобы в любой точке прямой L, параллельной V, она равнялась $P(L)$ *). Используя свойство инвариантности (24.61) с $\lambda = \bar{x}$, имеем при условии L

$$t(x) \mid L = \int_L f(x_1 - \bar{x},\ x_2 - \bar{x},\ \ldots,\ x_k - \bar{x})\,d\bar{x}.$$

Освобождаясь от условия, в качестве статистики критерия находим

$$t(x) = \int \left\{ \int_L f(x_1 - \bar{x},\ x_2 - \bar{x},\ \ldots,\ x_k - \bar{x})\,d\bar{x} \right\} dL,$$

откуда, заменяя $\bar{x}$ на u, т. е. интегрируя по L, получаем

$$t(x) = \int_{-\infty}^{\infty} f(x_1 - u,\ x_2 - u,\ \ldots,\ x_k - u)\,du. \qquad (24.67)$$

Критическая область размера α для этого несмещенного критерия определяется соотношением (24.62). Нетрудно видеть, что так полученный несмещенный критерий является единственным. Пример применения (24.67) дается в упражнении 24.15.

*) Следующий далее вывод формулы (24.67) неточен. Нетрудно видеть, что интеграл $P(L)$ по прямой L, параллельной вектору V и проходящей через точку $(x_1, \ldots, x_k)$, равен

$$P(L) = \sqrt{k} \int_{-\infty}^{\infty} f(x_1 - u,\ x_2 - u, \ldots,\ x_k - u)\,du.$$

Поскольку для статистики t постоянный множитель не играет роли, в (24.67) за статистику t принимается $P(L)/\sqrt{k}$ вместо $P(L)$. (*Прим. ред.*)

24.21 Обращаясь теперь к критериям, касающимся масштабных параметров, предположим, что совместное распределение k случайных величин имеет вид

$$dG = g\left(\frac{y_1}{\varphi_1}, \frac{y_2}{\varphi_2}, \ldots, \frac{y_k}{\varphi_k}\right)\frac{dy_1}{\varphi_1}\frac{dy_2}{\varphi_2}\ldots\frac{dy_k}{\varphi_k}, \qquad (24.68)$$

где все масштабные параметры φ_i положительны. Сделаем преобразование

$$x_i = \log|y_i|, \qquad \theta_i = \log\varphi_i$$

и найдем распределение x_i:

$$dF = g\{\exp(x_1-\theta_1), \exp(x_2-\theta_2), \ldots, \exp(x_k-\theta_k)\}\times$$
$$\times\exp\left\{\sum_{i=1}^{k}(x_i-\theta_i)\right\}dx_1\ldots dx_k. \qquad (24.69)$$

Распределение (24.69) имеет вид (24.59), рассмотренный выше. Проверка гипотезы

$$H_0'\colon \varphi_1 = \varphi_2 = \ldots = \varphi_k \qquad (24.70)$$

равносильна проверке H_0 из (24.60). Статистика (24.67) превращается здесь в

$$t(x) = \int_{-\infty}^{\infty} g\{\exp(x_1-u), \exp(x_2-u), \ldots, \exp(x_k-u)\}\times$$
$$\times\exp\left\{\sum_{i=1}^{k}x_i - ku\right\}du,$$

которая в терминах y_i переписывается как

$$t(y) = \prod_{i=1}^{k}|y_i|\int_0^{\infty} g\left(\frac{y_1}{v}, \frac{y_2}{v}, \ldots, \frac{y_k}{v}\right)\frac{dv}{v^{k+1}}. \qquad (24.71)$$

24.22 Рассмотрим теперь специальный случай k независимых случайных величин y_i/φ_i, имеющих гамма-распределения с параметрами m_i. Их совместное распределение есть

$$dG = \prod_{i=1}^{k}\left\{\frac{1}{\Gamma(m_i)}\left(\frac{y_i}{\varphi_i}\right)^{m_i-1}\right\}\exp\left(-\sum_{i=1}^{k}\frac{y_i}{\varphi_i}\right)\prod_{i=1}^{k}\frac{dy_i}{\varphi_i}. \qquad (24.72)$$

Для проверки гипотезы H_0' из (24.70) воспользуемся статистикой (24.71). Получаем

$$t(y) = \prod_{i=1}^{k}\left\{\frac{y_i^{m_i}}{\Gamma(m_i)}\right\}\int_0^{\infty}\exp\left(-\sum_{i=1}^{k}y_i/v\right)\frac{dv}{v^{m+1}}, \qquad (24.73)$$

где $m = \sum_{i=1}^{k} m_i$. Подставляя в (24.73) $u = \sum_{i=1}^{k} y_i/v$, находим

$$t(y) = \left[\frac{\Gamma(m)}{\prod_i \Gamma(m_i)}\right] \frac{\prod_i y_i^{m_i}}{\left(\sum_i y_i\right)^m}. \tag{24.74}$$

Отбросим постоянный множитель в квадратных скобках соотношения (24.74). Оставшаяся часть t достигает своего максимального значения T, когда $y_i \Big/ \sum_i y_i = m_i/m$, при этом

$$T = \prod_i m_i^{m_i}/m^m. \tag{24.75}$$

Обозначим теперь

$$t^* = -\log\left(\frac{t}{T}\right) = m \log\left(\frac{\sum_i y_i}{m}\right) - \sum_i m_i \log\left(\frac{y_i}{m_i}\right). \tag{24.76}$$

Статистика t^* дает несмещенный критерий для гипотезы H_0'. Она изменяется от 0 до ∞, и ее большие значения являются критическими.

Пример 24.6

Мы можем теперь применить (24.76) к задаче проверки гипотезы о равенстве дисперсий в k нормальных совокупностях, рассмотренной в примере 24.4. Действительно, при гипотезе H_0 каждая из величин $\sum_j (x_{ij} - \bar{x}_i)^2/(2\sigma^2)$ имеет гамма-распределение с параметром $\frac{1}{2}(n_i - 1)$. Подставляя, следовательно, в (24.76)

$$\left.\begin{aligned} y_i &= \sum (x_{ij} - \bar{x}_i)^2 = \nu_i s_i^2, \\ m_i &= \tfrac{1}{2}(n_i - 1) = \tfrac{1}{2}\nu_i, \\ m &= \sum_i m_i = \tfrac{1}{2}(n - k) = \tfrac{1}{2}\nu, \end{aligned}\right\} \tag{24.77}$$

находим в качестве статистики несмещенного критерия

$$2t^* = \nu \log\left(\frac{\sum_i \nu_i s_i^2}{\nu}\right) - \sum_i \nu_i \log(s_i^2). \tag{24.78}$$

Соотношение (24.78) идентично (24.44), так что $2t^*$ есть попросту статистика $-2 \log l^*$, которую мы там рассматривали. Таким

образом, как утверждалось в примере 24.4, критерий, основанный на l^*, является несмещенным. Отсюда видно, что неподправленная статистика ОП l из (24.40), в которой использовались другие веса, не может быть в общем случае несмещенной. Когда все выборки имеют одинаковые объемы, то, как показывает упражнение 24.7, эти два критерия эквивалентны. Даже в случае $k = 2$ при $n_1 \neq n_2$ неподправленный критерий ОП смещен (доказательство этого утверждения предоставлено читателю в упражнении 24.14).

24.23 Прежде чем закончить рассмотрение несмещенности критериев ОП, нужно упомянуть, что Полсон (1941) исследовал смещение некоторых критериев ОП для экспоненциальных распределений (некоторые из его результатов приведены в упражнениях 24.16 и 24.18). Кроме того, Дейли (1940) и Нарайн (1950) доказали несмещенность ряда критериев ОП для проверки независимости в многомерных нормальных совокупностях (к их результатам мы обратимся, когда встретимся с этими критериями в томе 3).

Другие свойства критериев ОП

24.24 Кроме состоятельности и несмещенности, что еще, в общем случае, можно сказать о свойствах критериев ОП? Прежде всего, мы знаем, что МП-оценки являются функциями от достаточных статистик (см. **18.4**) и, следовательно, статистика ОП (24.4) может быть переписана в виде

$$l = \frac{L(x \mid \theta_{r0}, t_s)}{L(x \mid T_{r+s})}, \tag{24.79}$$

где t_s — минимальный достаточный вектор для θ_s, когда выполняется гипотеза H_0, а T_{r+s} — достаточная статистика для всех параметров, когда H_0 не выполняется. Как мы видели в **17.38**, вообще говоря, неверно утверждать, что компоненты вектора T_{r+s} включают в себя компоненты t_s: достаточная статистика для θ_s при гипотезе H_0 может уже не составлять часть достаточной системы статистик, построенной, когда H_0 не выполняется, и даже если эта статистика входит в указанную систему, то она в отдельности может не быть достаточной для θ_s, будучи лишь частью вектора T_{r+s}, достаточного для (θ_r, θ_s). Таким образом, о статистике l можно лишь сказать, что она есть *некоторая* функция от двух систем достаточных статистик. В общем случае нет оснований предполагать, что она будет наиболее подходящей функцией.

Легко видеть, что метод ОП не обязательно порождает РНМ критерий, когда последний существует, так как даже в случае проверки простой гипотезы H_0 против простой альтернативы H_1 он не дает НКО (22.6).

Для отыскания РНМН критерия метод ОП непосредственно непригоден, поскольку он, вообще говоря, является смещенным,

но, как мы уже видели, с помощью простой поправки эта трудность может быть преодолена. Указанная поправка выражается в «перераспределении весов» в статистике критерия, что достигается заменой обычных МП-оценок несмещенными оценками (примеры 24.4—24.6). Иногда это эквивалентно исправлению критической области статистики, к которой приводит метод ОП (упражнение 24.14). Упражнение 24.16 показывает, что два РНМ критерия, выведенные в упражнениях 23.25, 23.26 для экспоненциального и прямоугольного распределений, получаются с помощью метода ОП, тогда как РНМН критерий для экспоненциального распределения, приведенный в упражнении 23.24, не эквивалентен критерию ОП, который является смещенным.

Вальд (1943а) доказал, что при условиях регулярности мощность критерия ОП имеет *асимптотически* некоторые оптимальные свойства (см., однако, **25.4** и пример 25.1). Гёфдинг (1965) получил оптимальное свойство критериев ОП для мультиномиальных распределений, когда размер $\alpha \to 0$ при возрастании объема выборки.

Принцип ОП является интуитивно привлекательным, когда нет «оптимального» критерия. Он имеет особенную ценность при проверке линейных гипотез (их мы будем рассматривать во второй части данной главы), где в общем случае не существует РНМН критерия. Но необходимо также напомнить, что в исключительных ситуациях метод ОП подвержен ошибкам. Следующий пример, принадлежащий Стейну и приведенный Леманом (1950), служит целебным предостережением против использования этого метода без исследования его свойств в каждом конкретном случае.

Пример 24.7

Дискретная случайная величина x принимает значения 0, ± 1, ± 2 со следующими вероятностями при гипотезе H_1:

$$\left.\begin{array}{lccccc} x: & 0 & \pm 1 & +2 & -2 \\ \mathbf{P}\,|\,H_1: & \alpha\left(\frac{1-\theta_1}{1-\alpha}\right) & \left(\frac{1}{2}-\alpha\right)\left(\frac{1-\theta_1}{1-\alpha}\right) & \theta_1\theta_2 & \theta_1(1-\theta_2) \end{array}\right\} \quad (24.80)$$

Параметры θ_1, θ_2 удовлетворяют неравенствам

$$0 \leqslant \theta_1 \leqslant \alpha < \frac{1}{2}, \qquad 0 \leqslant \theta_2 \leqslant 1,$$

где α — известная константа. Мы хотим на основе единственного наблюдения проверить простую гипотезу

$$H_0\colon\ \theta_1 = \alpha, \qquad \theta_2 = \frac{1}{2}$$

против общей альтернативы H_1 (24.80); при гипотезе H_0 вероятности имеют вид

$$\left.\begin{array}{lccccc} x: & 0 & \pm 1 & +2 & -2 \\ \mathsf{P}\,|\,H_0: & \alpha & \frac{1}{2}-\alpha & \frac{1}{2}\alpha & \frac{1}{2}\alpha \end{array}\right\}. \tag{24.81}$$

Если $x = 0, \pm 1$, то ФП не зависит от θ_2 и ее безусловный максимум достигается, когда θ_1 принимает минимально возможное значение, т. е. при $\hat{\theta}_1 = 0$. Следовательно, статистика ОП равна

$$l = \frac{L(x\,|\,H_0)}{L(x\,|\,\hat{\theta}_1, \hat{\theta}_2)} = 1 - \alpha, \quad x = 0, \pm 1. \tag{24.82}$$

Если $x = +2$ или -2, то ФП достигает безусловного максимума при выборе θ_2 соответственно максимально или минимально возможным, т. е. $\hat{\theta}_2 = 1$ или 0 соответственно, и при выборе θ_1 максимально возможным, т. е. $\hat{\theta}_1 = \alpha$. Максимальное значение ФП равно, следовательно, α, и статистика ОП имеет вид

$$l = \frac{1}{2}, \quad x = \pm 2. \tag{24.83}$$

Так как $\alpha < 1/2$, то из соотношений (24.82) и (24.83) следует, что критерий ОП заключается в отклонении гипотезы H_0, когда $x = \pm 2$. Из (24.81) видно, что размер данного критерия равен α. Но из (24.80) получаем, что его мощность равна в точности θ_1. Таким образом, для любого значения θ_1, удовлетворяющего неравенству

$$0 \leqslant \theta_1 < \alpha, \tag{24.84}$$

критерий ОП будет смещенным при всех θ_2, тогда как для $\theta_1 = \alpha$ при любых θ_2 он имеет мощность, равную его размеру α. В последнем случае приведенный критерий бесполезен, а в случае $0 \leqslant \theta_1 < \alpha$ он хуже, чем бесполезный, потому что критерий размера и мощности α для проверки гипотезы H_0 мы можем получить с помощью таблицы случайных чисел. Кроме того, существует полезный критерий, отвергающий H_0, когда $x = 0$. При этом по (24.81) его размер равен α, а мощность согласно (24.80) равна $\alpha\left(\frac{1-\theta_1}{1-\alpha}\right)$, что превосходит α, когда выполняется (24.84), и равно α, когда $\theta_1 = \alpha$.

Помимо дискретности случайной величины, примечательной особенностью этого предостерегающего примера является то, что интервал изменения одного из параметров определяется размером критерия α.

Д. Кокс (1961, 1962) рассматривал распределения статистик ОП, когда H_0 и H_1 образуют совершенно отдельные семейства сложных гипотез (так что (24.5) не выполняется), и получил некоторые результаты в случае больших выборок.

Общая линейная гипотеза и ее каноническая форма

24.25 Теперь мы в состоянии рассмотреть задачу проверки гипотез для общей линейной модели из главы 19. Как и в (19.8), положим

$$\boldsymbol{y} = \boldsymbol{X}\theta + \boldsymbol{\varepsilon}, \tag{24.85}$$

где ε_i некоррелированы, имеют нулевые средние и одинаковые дисперсии σ^2. Не будем пока делать дальнейших предположений о виде их распределений.

Предположим, что матрица $\boldsymbol{X}'\boldsymbol{X}$ невырождена (если это не так, то, как и в **19.13—16**, ее можно сделать невырожденный с помощью пополнения).

Пусть мы хотим проверить гипотезу

$$H_0\colon\ \boldsymbol{A}\theta = \boldsymbol{c}_0, \tag{24.86}$$

где $\boldsymbol{A}$ и $\boldsymbol{c}_0$ — известные $(r \times k)$-матрица и $(r \times 1)$-вектор. Гипотеза (24.86) накладывает $r (\leqslant k)$ ограничений, которые будем считать функционально независимыми, так что ранг матрицы $\boldsymbol{A}$ равен r. Гипотеза H_1 будет просто отрицанием H_0. Когда $r = k$, матрица $\boldsymbol{A}'\boldsymbol{A}$ невырождена и гипотеза (24.86) эквивалентна H_0: $\theta = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{c}_0$. Если $\boldsymbol{A}$ состоит из первых r строк $(n \times k)$-матрицы $\boldsymbol{X}$, то гипотеза H_0 имеет особенно простой смысл: это гипотеза о средних значениях первых r из величин y_i.

Рассмотрим $(n \times 1)$-вектор

$$\boldsymbol{z} = \boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{C}\hat{\theta}, \tag{24.87}$$

где $\boldsymbol{C}$ — матрица $(n \times k)$, а $\hat{\theta}$ — оценка наименьших квадратов (НК-оценка) для θ, полученная в (19.12). Тогда из (24.87) и (24.85) имеем

$$\boldsymbol{z} = \boldsymbol{C}\theta + \boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon},$$

так что

$$\mu = \mathsf{M}(\boldsymbol{z}) = \boldsymbol{C}\theta, \tag{24.88}$$

а матрица рассеяния вектора z, как и в **19.6**, равна

$$\boldsymbol{V} = \sigma^2\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{C}'. \tag{24.89}$$

Выберем теперь $\boldsymbol{C}$ так, чтобы компоненты вектора $\boldsymbol{z}$ были некоррелированы, т. е. чтобы

$$\boldsymbol{V} = \sigma^2\boldsymbol{I}.$$

В силу (24.89) для этого нужно, чтобы

$$\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{C}' = \boldsymbol{I}$$

или, в случае невырожденной матрицы $\boldsymbol{C}'\boldsymbol{C}$, чтобы

$$\boldsymbol{C}'\boldsymbol{C} = \boldsymbol{X}'\boldsymbol{X}. \tag{24.90}$$

Равенство (24.90) является условием некоррелированности величин z_i; из (24.90), (24.87) и (24.88) следует, что

$$(z-\mu)'(z-\mu)=\varepsilon' X(X'X)^{-1}X'\varepsilon=\{X(\hat{\theta}-\theta)\}'\{X(\hat{\theta}-\theta)\}. \quad (24.91)$$

24.26 Положим теперь

$$C=\begin{pmatrix} A \\ D \\ F \end{pmatrix},$$

где A есть $(r\times k)$-матрица из (24.86), D — некоторая $((k-r)\times k)$-матрица и F есть $((n-k)\times k)$-матрица, удовлетворяющая уравнению

$$F\theta=0. \quad (24.92)$$

Поскольку ранг матрицы A равен r, то можно D выбрать так, чтобы $(k\times k)$-матрица $\begin{pmatrix} A \\ D \end{pmatrix}$ была невырожденной и, таким образом, матрица C имела бы ранг k. Тогда ранг матрицы $C'C$ тоже равен k и, следовательно, она невырождена, как это требовалось выше, в (24.90).

Из (24.88) находим

$$\mu=\mathsf{M}(z)=\begin{pmatrix} A \\ D \\ F \end{pmatrix}\theta. \quad (24.93)$$

Следовательно, средние значения первых r величин z_i совпадают с левой частью (24.86); поэтому проверка гипотезы H_0 эквивалентна проверке

$$H_0:\ \mu_i=\mathsf{M}(z_i)=c_{0i}, \quad i=1,\ 2,\ \ldots,\ r, \quad (24.94)$$

— сложной гипотезы, накладывающей r ограничений на параметры. Так как согласно (24.92) последние $n-k$ значений μ_i равны нулю, то k параметров μ_i отличны от нуля, так что всего вместе с σ^2 имеется $k+1$ параметров.

24.27 Таким образом, мы привели задачу к следующему виду. Дано множество из n взаимно некоррелированных величин z_i с равными дисперсиями σ^2, причем $n-k$ величин z_i имеют нулевые средние, а у остальных k средние отличны от нуля. Гипотеза состоит в том, что r из этих k величин имеют заданные средние. Это так называемая каноническая форма общей линейной гипотезы.

Чтобы двигаться дальше, нам нужно сделать предположения о распределении ошибок в линейной модели (24.85). Будем считать, что каждая из величин ε_i распределена нормально, а вследствие их некоррелированности они независимы. Величины z_i,

являясь линейными функциями от них и будучи некоррелированными, также независимы и нормально распределены. Следовательно, из их совместного распределения находим ФП

$$L(\boldsymbol{z}\,|\boldsymbol{\mu},\ \sigma^2) = (2\pi\sigma^2)^{-n/2}\exp\Big\{-\frac{1}{2\sigma^2}(\boldsymbol{z}-\boldsymbol{\mu})'(\boldsymbol{z}-\boldsymbol{\mu})\Big\} =$$

$$= (2\pi\sigma^2)^{-n/2}\exp\Big[-\frac{1}{2\sigma^2}\{(\boldsymbol{z}_r-\boldsymbol{\mu}_r)'(\boldsymbol{z}_r-\boldsymbol{\mu}_r)+$$

$$+(\boldsymbol{z}_{k-r}-\boldsymbol{\mu}_{k-r})'(\boldsymbol{z}_{k-r}-\boldsymbol{\mu}_{k-r})+\boldsymbol{z}'_{n-k}\boldsymbol{z}_{n-k}\}\Big], \qquad (24.95)$$

где индексы у векторов указывают их размерность. Наша гипотеза есть

$$H_0\colon\ \boldsymbol{\mu}_r = \boldsymbol{c}_0, \qquad (24.96)$$

а H_1 представляет собой ее отрицание.

В примере 23.14 мы видели, что если имеется только одно ограничение ($r = 1$), то существует РНМН критерий для гипотезы H_0 против H_1. Это очевидно также из того факта, то в данном случае проверяется гипотеза о среднем одномерной нормальной совокупности с неизвестной дисперсией: РНМН критерием для этой гипотезы (как видно из примера 23.14) служит обычный t-критерий Стьюдента с равными «хвостами».

Колоджейчик (1935), которому принадлежат первые общие результаты о линейных гипотезах, доказал отсутствие РНМ критерия в ситуации с более чем одним ограничением и показал, что имеется пара подобных односторонних РНМ критериев, когда $r = 1$: это односторонние t-критерии Стьюдента (см. упражнение 23.7). Только что было указано на существование двустороннего t-критерия Стьюдента, являющегося РНМН в случае $r = 1$, но критическая область этого критерия зависит от того, о каком из μ_i проверяется гипотеза. Таким образом, при $r > 1$ нет общей РНМН критической области.

Так как нельзя найти «оптимального» в каком-либо смысле критерия, то мы обратимся к методу ОП, позволяющему получить интуитивно разумный критерий.

24.28 Вывод статистики ОП достаточно прост. Безусловный максимум выражения (24.95) определяется с помощью решения системы уравнений

$$\frac{\partial \log L}{\partial \mu_i} = 0, \quad i = 1, 2, \ldots, k, \quad \frac{\partial \log L}{\partial (\sigma^2)} = 0.$$

Это дает МП-оценки

$$\hat{\mu}_i = z_i, \quad i = 1, 2, \ldots, k, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=k+1}^{n} z_i^2,$$

откуда

$$(\boldsymbol{z}-\hat{\boldsymbol{\mu}})'(\boldsymbol{z}-\hat{\boldsymbol{\mu}}) = n\hat{\sigma}^2.$$

Итак, находим безусловный максимум ФП

$$L(\boldsymbol{z}\,|\,\hat{\boldsymbol{\mu}},\ \hat{\sigma}^2) = (2\pi\hat{\sigma}^2 e)^{-n/2} = \left(\frac{2\pi e}{n}\sum_{i=k+1}^{n} z_i^2\right)^{-n/2} \tag{24.97}$$

Когда выполняется гипотеза (24.96), имеем МП-оценки для мешающих параметров

$$\hat{\hat{\mu}}_i = z_i, \quad i = r+1, \quad r+2, \ \ldots, \ k,$$

$$\hat{\hat{\sigma}}^2 = \frac{1}{n}\left\{\sum_{i=k+1}^{n} z_i^2 + \sum_{i=1}^{r}(z - c_{0i})^2\right\},$$

откуда

$$(\boldsymbol{z} - \hat{\hat{\boldsymbol{\mu}}})'(\boldsymbol{z} - \hat{\hat{\boldsymbol{\mu}}}) = n\hat{\hat{\sigma}}^2,$$

так что условный максимум ФП равен

$$L(\boldsymbol{z}\,|\,\boldsymbol{c}_0,\ \hat{\hat{\boldsymbol{\mu}}},\ \hat{\hat{\sigma}}^2) = (2\pi\hat{\hat{\sigma}}^2 e)^{-n/2} =$$

$$= \left[\frac{2\pi e}{n}\left\{\sum_{i=k+1}^{n} z_i^2 + \sum_{i=1}^{r}(z_i - c_{0i})^2\right\}\right]^{-n/2}. \tag{24.98}$$

Из соотношений (24.97) и (24.98) находим статистику ОП l:

$$l^{2/n} = \frac{\hat{\sigma}^2}{\hat{\hat{\sigma}}^2} = \frac{1}{1+W}, \tag{24.99}$$

где

$$W = (\boldsymbol{z}_r - \boldsymbol{c}_0)'(\boldsymbol{z}_r - \boldsymbol{c}_0)/\boldsymbol{z}'_{n-k}\boldsymbol{z}_{n-k} =$$

$$= \frac{\sum_{i=1}^{r}(z_i - c_{0i})^2}{\sum_{i=k+1}^{n} z_i^2} = \frac{\hat{\hat{\sigma}}^2 - \hat{\sigma}^2}{\hat{\sigma}^2}. \tag{24.100}$$

Заметим, что значения $n\hat{\hat{\sigma}}^2$ и $n\hat{\sigma}^2$ являются минимумами по $\boldsymbol{\mu}$ функции $(\boldsymbol{z} - \boldsymbol{\mu})'(\boldsymbol{z} - \boldsymbol{\mu})$ при гипотезах H_0 и H_1 соответственно. Согласно (24.91) это то же самое, что минимумы функции

$$R = \{\boldsymbol{X}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\}'\{\boldsymbol{X}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\}$$

относительно $\boldsymbol{\theta}$. Тождество по $\boldsymbol{\theta}$

$$S = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\theta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\theta}) \equiv (\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\theta}})'(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\theta}}) + R$$

легко проверить непосредственно. При этом слагаемое справа

$$(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\theta}})'(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\theta}})$$

не зависит от $\boldsymbol{\theta}$. Минимизация функции R по $\boldsymbol{\theta}$ эквивалентна, таким образом, минимизации функции S. Но процесс минимизации S относительно $\boldsymbol{\theta}$ есть в точности способ, с помощью которого мы пришли к оценкам наименьших квадратов в **19.4**. Следовательно, для получения $\hat{\hat{\sigma}}^2$ и $\hat{\sigma}^2$ в соотношении (24.100) мы минимизируем функцию S в исходной модели при гипотезах H_0 и H_1 соответственно.

Поскольку l есть монотонно убывающая функция от W, то критерий ОП эквивалентен тому, чтобы отвергать гипотезу H_0, когда значение W велико. Если разделить числитель и знаменатель величины W на σ^2, то при гипотезе H_0 W превращается в отношение суммы квадратов r независимых нормальных величин к независимой сумме квадратов $n-k$ таких же величин, т. е. является отношением двух независимых χ^2-величин с r и $n-k$ степенями свободы. Итак, когда выполняется гипотеза H_0, статистика $F = \frac{(n-k)}{r} W$ распределена, как дисперсионное отношение с $(r, n-k)$ степенями свободы (см. **16.15**), и критерий ОП формулируется в терминах величины F, причем критическая область образована ее большими значениями.

Многие стандартные статистические критерии могут быть сведены к критериям для линейных гипотез, и мы часто будем встречаться с ними в последующих главах.

Пример 24.8

Как специальный, особенно важный, случай рассмотрим гипотезу

$$H_0\colon\ \boldsymbol{\theta}_r = \mathbf{0},$$

где $\boldsymbol{\theta}_r$ есть $(r \times 1)$-подвектор вектора $\boldsymbol{\theta}$ в (24.85). Мы можем, следовательно, переписать соотношение (24.85) в виде

$$\boldsymbol{y} = (\boldsymbol{X}_1 \boldsymbol{X}_2)\begin{pmatrix} \boldsymbol{\theta}_r \\ \boldsymbol{\theta}_{k-r} \end{pmatrix} + \boldsymbol{\varepsilon},$$

где $\boldsymbol{X}_1$ и $\boldsymbol{X}_2$ — матрицы порядка $(n \times r)$ и $(n \times (k-r))$ соответственно. Тогда гипотеза H_0 эквивалентна предположению, что

$$\boldsymbol{y} = \boldsymbol{X}_2 \boldsymbol{\theta}_{k-r} + \boldsymbol{\varepsilon}.$$

Согласно **24.28** мы ищем минимумы функции $S = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}$. Так как и при H_0 и при H_1 оцениваются все параметры некоторой линейной модели, то можно воспользоваться результатом из **19.9**. Мы получаем, что минимум при гипотезе H_0 равен

$$n\hat{\hat{\sigma}}^2 = \boldsymbol{y}' \left\{ \boldsymbol{I} - \boldsymbol{X}_2 (\boldsymbol{X}_2' \boldsymbol{X}_2)^{-1} \boldsymbol{X}_2' \right\} \boldsymbol{y},$$

тогда как при H_1 он равен

$$n\hat{\sigma}^2 = \boldsymbol{y}'\{\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\}\boldsymbol{y},$$

где $\boldsymbol{X} = (\boldsymbol{X}_1\boldsymbol{X}_2)$. Статистика

$$F = \frac{n-k}{r}\left(\frac{\hat{\hat{\sigma}}^2 - \hat{\sigma}^2}{\hat{\sigma}^2}\right)$$

распределена, как дисперсионное отношение с $(r, n-k)$ степенями свободы, и критическая область для H_0 образована 100α процентами наибольших значений величины F.

24.29 В **24.28** мы видели, что критерий ОП основан на статистике (24.100), которую можно переписать в виде

$$W = \frac{\sum_{i=1}^{r}(z_i - c_{0i})^2/\sigma^2}{\sum_{i=k+1}^{n} z_i^2/\sigma^2}.$$

Независимо от того, верна гипотеза H_0 или нет, знаменатель величины W имеет χ^2-распределение с $n-k$ степенями свободы. Когда H_0 выполняется, то, как мы видели, числитель является также χ^2-величиной с r степенями свободы. Если же H_0 не выполняется, то последнее утверждение неверно. В действительности числитель будет всегда нецентральной χ^2-величиной (см. **24.4**) с r степенями свободы и параметром нецентральности

$$\lambda = \sum_{i=1}^{r}(c_{0i} - \mu_i)^2/\sigma^2 = (\boldsymbol{\mu}_r - \boldsymbol{c}_0)'(\boldsymbol{\mu}_r - \boldsymbol{c}_0)/\sigma^2, \qquad (24.101)$$

где μ_i — истинное среднее величины z_i. Только когда имеет место гипотеза H_0, λ равно нулю, и получается центральное χ^2-распределение, как в **24.28**. Так как для нахождения мощности критерия ОП нам требуется распределение величины W (или эквивалентной ей величины F), когда H_0 не выполнена, то мы приходим к необходимости изучения отношения нецентральной χ^2-величины к центральной.

Нецентральное F-распределение

24.30 Рассмотрим вначале отношение двух независимых величин z_1, z_2, имеющих нецентральные χ^2-распределения (24.18) со степенями свободы ν_1, ν_2 и параметрами нецентральности λ_1, λ_2 соответственно. Используя (11.74), находим распределение

для $u = z_1/z_2$:

$$dH(u) =$$

$$= du \int_0^\infty \frac{e^{-\frac{1}{2}(uv+\lambda_1)}(uv)^{\frac{1}{2}\nu_1-1}}{2^{\nu_1/2}\Gamma\left\{\frac{1}{2}(\nu_1-1)\right\}\Gamma\left(\frac{1}{2}\right)} \sum_{r=0}^\infty \frac{\lambda_1^r (uv)^r}{(2r)!} \mathrm{B}\left\{\frac{1}{2}(\nu_1-1),\ \frac{1}{2}+r\right\} \times$$

$$\times \frac{e^{-\frac{1}{2}(v+\lambda_2)} v^{\frac{1}{2}\nu_2-1}}{2^{\nu_2/2}\Gamma\left\{\frac{1}{2}(\nu_2-1)\right\}\Gamma\left(\frac{1}{2}\right)} \sum_{s=0}^\infty \frac{\lambda_2^s v^s}{(2s)!} \mathrm{B}\left\{\frac{1}{2}(\nu_2-1),\ \frac{1}{2}+s\right\} v\,dv.$$

Обозначая $\lambda = \lambda_1 + \lambda_2$, $\nu = \nu_1 + \nu_2$, после упрощений получим

$$dH(u) =$$

$$= \frac{e^{-\frac{1}{2}\lambda}}{2^{\nu/2}} \sum_{r=0}^\infty \sum_{s=0}^\infty \frac{\lambda_1^r}{(2r)!} \frac{\lambda_2^s}{(2s)!} \frac{\Gamma\left(\frac{1}{2}+r\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\nu_1+r\right)} \cdot \frac{\Gamma\left(\frac{1}{2}+s\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\nu_2+s\right)} \times$$

$$\times \left\{\int_0^\infty e^{-\frac{1}{2}v(1+u)} v^{\frac{1}{2}\nu+r+s-1}\,dv\right\} u^{\frac{1}{2}\nu_1+r-1}\,du. \qquad (24.102)$$

Интеграл в (24.102) равен $\Gamma\left(\frac{1}{2}\nu+r+s\right)\Big/\left(\frac{1+u}{2}\right)^{\frac{1}{2}\nu+r+s}$. Таким образом,

$$dH(u) = e^{-\frac{1}{2}\lambda} \sum_{r=0}^\infty \sum_{s=0}^\infty \frac{\lambda_1^r \lambda_2^s}{(2r)!\,(2s)!} \frac{\Gamma\left(\frac{1}{2}+r\right)\Gamma\left(\frac{1}{2}+s\right)2^{r+s}}{\mathrm{B}\left(\frac{1}{2}\nu_1+r,\ \frac{1}{2}\nu_2+s\right)\left\{\Gamma\left(\frac{1}{2}\right)\right\}^2} \times$$

$$\times u^{\frac{1}{2}\nu_1+r-1}\left(\frac{1}{1+u}\right)^{\frac{1}{2}\nu+r+s} du. \qquad (24.103)$$

Поскольку

$$\frac{\Gamma\left(\frac{1}{2}+r\right)\Gamma\left(\frac{1}{2}+s\right)2^{r+s}}{(2r)!\,(2s)!\left\{\Gamma\left(\frac{1}{2}\right)\right\}^2} = \frac{1}{2^{r+s}r!\,s!},$$

то (24.103) окончательно приводится к виду

$$dH(u) = e^{-\frac{1}{2}\lambda} \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{(\lambda_1/2)^r}{r!} \frac{(\lambda_2/2)^s}{s!} u^{\frac{1}{2}\nu_1+r-1} \times$$

$$\times \left(\frac{1}{1+u}\right)^{\frac{1}{2}\nu+r+s} \frac{du}{\mathrm{B}\left(\frac{1}{2}\nu_1+r, \frac{1}{2}\nu_2+s\right)}. \qquad (24.104)$$

Этот результат получил Тэнг (1938) и изучал Прайс (1964).

24.31 Для того чтобы найти распределение отношения нецентральной χ^2-величины, деленной на число ее степеней свободы, к независимой от нее центральной χ^2-величине, тоже поделенной на число степеней свободы, положим в (24.104)

$$F' = \frac{z_1/\nu_1}{z_2/\nu_2} = \frac{\nu_2}{\nu_1} u, \qquad \lambda_2 = 0, \qquad \lambda_1 = \lambda.$$

Тогда

$$dG(F') =$$

$$= e^{-\frac{1}{2}\lambda} \sum_{r=0}^{\infty} \frac{(\lambda/2)^r (\nu_1/\nu_2)^{\frac{1}{2}\nu_1+r}}{r!\,\mathrm{B}\left(\frac{1}{2}\nu_1+r, \frac{1}{2}\nu_2\right)} \cdot \frac{(F')^{\frac{1}{2}\nu_1+r-1}}{\left(1+\frac{\nu_1}{\nu_2}F'\right)^{\frac{1}{2}(\nu_1+\nu_2)+r}} dF'. \qquad (24.105)$$

Выражение (24.105) является обобщением распределения дисперсионного отношения (16.24) (F-распределения), к которому оно сводится, когда $\lambda = 0$. Это так называемое нецентральное F-распределение со степенями свободы ν_1, ν_2 и параметром нецентральности λ. Мы будем иногда обозначать его $F'(\nu_1, \nu_2, \lambda)$. Подобно распределению (24.18), оно впервые рассматривалось Фишером (1928а) и было изучено Уишартом (1932), Тэнгом (1938), Патнайком (1949), Прайсом (1964) и Тайкью (1965а).

Функция мощности критерия ОП для линейной гипотезы

24.32 Из **24.28—29** и **24.31** сразу следует, что функция мощности критерия ОП для общей линейной гипотезы имеет вид

$$P = \int_{F'_\alpha(\nu_1, \nu_2, 0)}^{\infty} dG\{F'(\nu_1, \nu_2, \lambda)\}, \qquad (24.106)$$

где F'_α есть $100(1-\alpha)$-процентная точка распределения, $\nu_1 = r$, $\nu_2 = n - k$, а параметр λ определен в (24.101).

Для P были построены различные таблицы и графики:

(1) Тэнг (1938) приводит дополнение к величине P (т. е. ошибку II рода) для размеров критерия $\alpha = 0{,}01,\ 0{,}05$; ν_1 (у него f_1) $= 1(1)8$, ν_2 (у него f_2) $= 2(2)6(1)30,\ 60,\ \infty$ и $\varphi = \{\lambda/(\nu_1+1)\}^{1/2} = 1(0{,}5)3(1)8$. Эти таблицы воспроизведены Манном (1949) и Кемпторном (1952).

(2) Лемер (1944) приводит обратные таблицы для φ при $\alpha = 0{,}01,\ 0{,}05$; $\nu_1 = 1(1)10,\ 12,\ 15,\ 20,\ 24,\ 30,\ 40,\ 60,\ 80,\ 120,\ \infty$; $\nu_2 = 2(2)20,\ 24,\ 30,\ 40,\ 60,\ 80,\ 120,\ 240,\ \infty$ и P (у нее β) $= 0{,}7,\ 0{,}8$.

(3) Э. Пирсон и Хартли (1951) приводят восемь графиков функции мощности, соответствующих значениям ν_1 от 1 до 8. Каждый график показывает значения мощности для $\nu_2 = 6(1)10,\ 12,\ 15,\ 20,\ 30,\ 60,\ \infty$; $\alpha = 0{,}01,\ 0{,}05$ и для φ, изменяющейся от 1 (за исключением случая $\nu_1 = 1$, когда φ изменяется от 2 для $\alpha = 0{,}01$ и от 1,2 для 0,05) до значения, достаточного для того, чтобы мощность равнялась по крайней мере 0,98. Для $\nu_1 = 1$ этот график воспроизведен в *Biometrika Tables* *).

4) М. Фокс (1956) приводит обратные графики для каждой из комбинаций $\alpha = 0{,}01,\ 0{,}05$ с мощностью P (у него β) $= 0{,}5,\ 0{,}7,\ 0{,}8,\ 0{,}9$. Каждый график дает линии уровня функции φ для $\nu_1 = 3(1)10(2)20(20)100,\ 200,\ \infty$; $\nu_2 = 4(1)10(2)20(20)100,\ 200,\ \infty$. Приводится также номограмма для каждого α, облегчающая интерполяцию по β.

(5) Дункан (1957) приводит графики для $\alpha = 0{,}01$ и $\alpha = 0{,}05$, показывающие для $\nu_2 = 6(1)10,\ 12,\ 15,\ 20(10)40,\ 60,\ \infty$ значения ν_1 (изменяющиеся от 1 до 8) и φ, требуемые для того, чтобы получить мощность $P = 1 - \beta = 0{,}50$ и $0{,}90$.

Аппроксимация функции мощности критерия ОП

24.33 Как видно из формы распределения (24.105), вычисление точной функции мощности (24.106) является трудоемким делом, и даже теперь она табулирована далеко не полностью. Однако можно получить простую аппроксимацию функции мощности, пользуясь методом и результатами нашей аппроксимации нецентрального χ^2-распределения в **24.5**. Если z_1 есть нецентральная χ^2-величина с ν_1 степенями свободы и параметром нецентральности λ, то из соотношения (24.21) получаем, что $z_1 \Big/ \left(\frac{\nu_1 + 2\lambda}{\nu_1 + \lambda}\right)$ является приближенно центральной χ^2-величиной, имеющей $\nu^* = (\nu_1 + \lambda)^2/(\nu_1 + 2\lambda)$ степеней свободы. Таким образом,

$$z_1 \Big/ \left\{ \nu^* \left(\frac{\nu_1 + 2\lambda}{\nu_1 + \lambda} \right) \right\} = z_1/(\nu_1 + \lambda)$$

есть приближенно центральная χ^2-величина, деленная на число ее степеней свободы ν^*. Следовательно, z_1/ν_1 равна приближенно этой величине, умноженной на $(\nu_1 + \lambda)/\nu_1$.

Если определим теперь нецентральную F'-величину, полагая

$$F' = \frac{z_1/\nu_1}{z_2/\nu_2},$$

---

*) Графики (3) воспроизведены в книге *Таблицы математической статистики*; графики (4) — в книге *Сборник статистических таблиц*. (*Прим. ред.*)

где z_2 — центральная χ^2-величина с ν_2 степенями свободы, то из предыдущего сразу следует, что приближенно

$$F' = \frac{\nu_1 + \lambda}{\nu_1} F, \tag{24.107}$$

где F — центральная F-величина со степенями свободы $\nu^* = (\nu_1 + \lambda)^2/(\nu_1 + 2\lambda)$ и ν_2.

Простая аппроксимация (24.107) оказывается удивительно эффективной. Патнайк, сравнивая значения функции мощности, взятые из точных таблиц Тэнга (1938), с вычисленными по приближению (24.107), показал, что у них совпадают первые две значащие цифры. В практических задачах этого достаточно.

Таким образом, чтобы вычислить мощность критерия ОП для линейной гипотезы, можно аппроксимировать (24.106) центральным F-интегралом

$$P = \int_{\left(\frac{\nu_1}{\nu_1+\lambda}\right) F_\alpha(\nu_1, \nu_2)}^{\infty} dG\left\{ F\left(\frac{(\nu_1+\lambda)^2}{(\nu_1+2\lambda)}, \nu_2\right)\right\}, \tag{24.108}$$

причем размер критерия определяется, если положить $\lambda = 0$. Величина $(\nu_1 + \lambda)^2/(\nu_1 + 2\lambda)$ является в общем случае дробной, и поэтому необходима интерполяция. Однако даже центральное F-распределение еще не столь хорошо табулировано, чтобы можно было легко найти точное значение интеграла (24.108) (см. список таблиц в **16.19**).

Принадлежащая Тайкью (1965а, 1966) аппроксимация центральным F распределением по трем моментам является еще более точной.

Дар (1962) дает простую нормальную аппроксимацию для распределения отношения двух независимых одинаково распределенных нецентральных F-величин.

Нецентральное t-распределение

24.34 При $\nu_1 = 1$ нецентральное F-распределение (24.105) сводится к нецентральному t^2-распределению, в точности как для центральных распределений (см. **16.15**). Если мы перейдем от t^2 к t, то получим нецентральное t-распределение, которое будем называть t'-распределением. Структура нецентральной χ^2-величины, как суммы квадратов нецентральных нормальных величин, показывает, что можно записать

$$t' = (z + \delta)/w^{1/2}, \tag{24.109}$$

где z — нормальная величина со средним нуль, а w — независимая от нее χ^2/f-величина с f степенями свободы (в данном случае мы пишем f вместо ν_2 в (24.105) и $\delta^2 = \lambda$).

Все сказанное о F'-распределении относится и к t'^2-распределению. Однако распределение t' привлекло к себе особое внимание вследствие его важности в приложениях.

Джонсон и Уэлч (1939) изучали t'-распределение и вычислили таблицы для его $100(1-\alpha)$-процентных точек при α или $1-\alpha = 0{,}005$, 0,01, 0,025, 0,05, 0,1(0,1)0,5; $f = 4(1)9$, 16, 36, 144, ∞ и любых δ и, наоборот, для нахождения δ при заданных значениях t'. Резников (1962) приводит дополнительные таблицы.

Резников и Либерман (1957) дают таблицы плотности и функции распределения величины t' с 4 десятичными знаками с шагом по $t/f^{1/2}$, равным 0,05, при $f = 2(1)24(5)49$ и для значений δ, определяемых из уравнения

$$\int_{\delta/(f+1)^{1/2}}^{\infty} (2\pi)^{-1/2} \exp\left(-\frac{1}{2}x^2\right) dx = \alpha,$$

с $\alpha = 0{,}001$, 0,0025, 0,004, 0,01, 0,025, 0,04, 0,065, 0,10, 0,15, 0,25. Они, а также Шойер и Спёрджен (1963) вычислили некоторые процентные точки этих распределений. Локс и др. (1963) приводят аналогичные таблицы с шагом по t', равным 0,2, для $f = 1(1)20(5)40$ и δ, определяемых из уравнения $\delta(f+1)^{-1/2}$ или $\delta(f+2)^{-1/2} = 0(0{,}25)3$. Обширные таблицы процентных точек приводит Оуэн (1963)*). Хогбен и др. (1961) дали метод получения моментов и таблицы для первых четырех моментов. Эймос (1964) изучал аппроксимации этого распределения с помощью рядов.

Кришнан (1967) изучал моменты величины (24.109), полученной из (24.104), а не из (24.105) (так что w являлась нецентральной χ^2-величиной), и аппроксимировал ее распределение.

24.35 Особенно важным приложением t'-распределения является нахождение мощности t-критерия Стьюдента, критическая область которого состоит только из одного «хвоста» (случай равных «хвостов» соответствует, разумеется, t'^2-распределению). Критерий служит для проверки гипотезы, что в (24.109) $\delta = 0$, его критическая область определяется с помощью центрального t-распределения. Его мощность, очевидно, равна интегралу от нецентрального t-распределения по критической области. Она была специально табулирована Нейманом и др. (1935), которые для $\alpha = 0{,}05$, 0,01; f (у них n) $= 1(1)30$, ∞ и δ (у них ρ) $= 1(1)10$ приводят таблицы и графики для дополнения мощности $1-P$, а также значения δ, для которых $P = 1-\alpha$. Нейман и Токарска (1936) вычислили обратные таблицы величины δ для тех же значений α, f и $1-P = 0{,}05$, 0,10(0,10)0,90. Оуэн (1965) приводит таблицы для δ с 5 десятичными знаками при $\alpha = 0{,}05$, 0,025, 0,01 и 0,005; $(n-1) = 1(1)30(5)100(10)200$, ∞ и $1-P = 0{,}01$, 0,05, 0,10(0,10)0,90.

Оптимальные свойства критерия ОП для общей линейной гипотезы

24.36 В **24.27** мы видели, что (кроме случая $r = 1$) для общей линейной гипотезы не существует РНМН критерия. Тем не менее критерий ОП для этой гипотезы обладает определенными оптимальными свойствами, к изложению которых мы сейчас перехо-

*) См. *Сборник статистических таблиц.* (*Прим. ред.*)

дим. При этом будут использованы упрощенные доказательства, принадлежащие Волфовицу (1949) и Леману (1950).

В **24.28** были получены МП-оценки для $k-r+1$ мешающих параметров при гипотезе H_0. Они являются компонентами вектора

$$t=(\hat{\hat{\mu}},\ \hat{\hat{\sigma}}^2),$$

определенного перед (24.98). Когда выполнена гипотеза H_0, компоненты вектора t образуют систему из $k-r+1$ достаточных статистик для мешающих параметров. Согласно **23.10** их распределение полно. Таким образом, как показано в **23.19**, каждая подобная критическая область w размера α для H_0 будет содержать долю α каждой поверхности $t=\text{const}$. На этой поверхности каждая компонента вектора t, в частности и $\hat{\hat{\sigma}}^2$, постоянна. Положим

$$n\hat{\hat{\sigma}}^2=\sum_{i=k+1}^{n} z_i^2+\sum_{i=1}^{r}(z_i-c_{0i})^2=a^2, \tag{24.110}$$

где a — константа.

Рассмотрим теперь фиксированное значение величины λ, определенной в (24.101), например $\lambda=d^2>0$. Мощность любой подобной области на этой поверхности будет складываться из мощностей на (24.110) при всех a. Для фиксированного a мощность на поверхности $\lambda=d^2$ равна

$$\mathbf{P}(w\,|\,\lambda,\ a)=\int\limits_{\lambda=d^2} L(\boldsymbol{z}\,|\,\boldsymbol{\mu},\ \sigma^2)\,d\boldsymbol{z}, \tag{24.111}$$

где L есть ФП, определенная в (24.95). Это можно переписать подробно в виде

$$\mathbf{P}(w\,|\,\lambda,\ a)=$$

$$=(2\pi\sigma^2)^{-n/2}\int\limits_{\lambda=d^2}\exp\left\{-\frac{1}{2\sigma^2}[\{(\boldsymbol{z}_r-\boldsymbol{c}_0)-(\boldsymbol{\mu}_r-\boldsymbol{c}_0)\}'\{(\boldsymbol{z}_r-\boldsymbol{c}_0)-\right.$$

$$\left.-(\boldsymbol{\mu}_r-\boldsymbol{c}_0)\}+(\boldsymbol{z}_{k-r}-\boldsymbol{\mu}_{k-r})'(\boldsymbol{z}_{k-r}-\boldsymbol{\mu}_{k-r})+\boldsymbol{z}'_{n-k}\boldsymbol{z}_{n-k}]\right\}d\boldsymbol{z}. \tag{24.112}$$

Используя (24.110) и (24.101), из (24.112) находим

$$\mathbf{P}(w\,|\,\lambda,\ a)=$$

$$=(2\pi\sigma^2)^{-(n-k+r)/2}\exp\left\{-\frac{1}{2}\left(d^2+\frac{a^2}{\sigma^2}\right)\right\}\int\limits_{\lambda=d^2}\exp\{(\boldsymbol{z}_r-\boldsymbol{c}_0)'(\boldsymbol{\mu}_r-\boldsymbol{c}_0)\}\,d\boldsymbol{z}_r. \tag{24.113}$$

Вектор $\boldsymbol{z}_{k-r}$ исключен в результате интегрирования по всей области его изменения, так как его распределение не зависит от λ и a. В соотношении (24.113) непостоянным множителем является

только интеграл. Его нужно сделать максимальным для того, чтобы получить критическую область w с максимальным P. Интеграл берется по поверхности $\lambda = d^2$ или $(\mu_r - c_0)'(\mu_r - c_0) = \text{const}$. Он, очевидно, представляет собой монотонно возрастающую функцию от $|z_r - c_0|$, т. е. от

$$(z_r - c_0)'(z_r - c_0) = \sum_{i=1}^{r} (z_i - c_{0i})^2.$$

Если при фиксированном a в (24.110) $\sum_{i=1}^{r} (z_i - c_{0i})^2$ является максимальной, то величина W, определенная в (24.100), тоже максимальна. Таким образом, при фиксированных λ и a максимальное значение величины $\mathbf{P}(w|\lambda, a)$ достигается, когда w состоит из больших значений W. Поскольку приведенное утверждение выполняется для каждого a, то оно будет иметь место и без предположения, что a фиксировано. Следовательно, мы установили, что на любой поверхности $\lambda = d^2 > 0$ критерий ОП, отвергающий гипотезу при больших значениях W, имеет максимальную мощность. Полученный результат принадлежит Вальду (1942).

В качестве непосредственного следствия получается результат Сюя (1941), утверждающий, что критерий ОП является РНМ среди всех критериев, функции мощности которых зависят только от λ.

Инвариантные критерии

24.37 Рассматривая несмещенные критерии для параметров сдвига в **24.20**, мы нашли совершенно естественным ввести условие инвариантности (24.61), как необходимое условие, которому должен удовлетворять любой приемлемый критерий. Аналогично, для масштабного параметра в **24.21**, переходя с помощью логарифмического преобразования от (24.68) к (24.69), мы неявно предполагали, что статистика t должна удовлетворять условию

$$t(y_1, y_2, \ldots, y_n) = t(cy_1, cy_2, \ldots, cy_n), \quad c > 0. \quad (24.114)$$

Часто бывает разумным ограничить класс рассматриваемых критериев теми критериями, которые инвариантны относительно преобразований, оставляющих инвариантной проверяемую гипотезу. Если этого не сделать, например, при проверке равенства параметров сдвига (или масштаба), то изменение начала координат (или единицы измерения) будет влиять на заключения, к которым приводит критерий. Связь между принципами инвариантности и достаточности в общем случае обсуждалась Холлом и др. (1965). Они приводят теорему, принадлежащую Стейну

и дающую условия, при которых не имеет значения порядок применения этих принципов.

Если рассмотреть с этой точки зрения каноническую форму общей линейной гипотезы из **24.27**, то сразу видно, что задача является инвариантной при

(а) любом ортогональном преобразовании вектора $(\boldsymbol{z}_r - \boldsymbol{c}_0)$ (оно оставляет неизменным величину $(\boldsymbol{z}_r - \boldsymbol{c}_0)'(\boldsymbol{z}_r - \boldsymbol{c}_0)$);

(б) любом ортогональном преобразовании вектора $\boldsymbol{z}_{n-k}$ (оно не изменяет $\boldsymbol{z}'_{n-k}\boldsymbol{z}_{n-k}$);

(в) прибавлении любой константы a к каждой компоненте вектора $\boldsymbol{z}_{n-k}$ (у которого вектор средних произволен);

(г) умножении всех величин на $c > 0$ (что влияет только на общую дисперсию σ^2).

Легко видеть, что статистика t инвариантна при всех преобразованиях (а)—(г) тогда и только тогда, когда она есть функция только от $W = (\boldsymbol{z}_r - \boldsymbol{c}_0)'(\boldsymbol{z}_r - \boldsymbol{c}_0)/\boldsymbol{z}'_{n-k}\boldsymbol{z}_{n-k}$. Ясно, что если t зависит лишь от W, то ее функция мощности, как и функция мощности W, будет зависеть только от λ. Следовательно, согласно последнему предложению из **24.36** критерий ОП, отвергающий гипотезу при больших значениях W, является РНМ среди всех инвариантных критериев для общей линейной гипотезы.

УПРАЖНЕНИЯ

24.1 Показать, что х. ф. нецентрального χ^2-распределения (24.18) равна

$$\varphi(t) = (1 - 2it)^{-\nu/2} \exp\left\{\frac{\lambda it}{1 - 2it}\right\},$$

откуда получаются семиинварианты $\varkappa_r = (\nu + r\lambda)2^{r-1}(r-1)!$. В частности,

$$\varkappa_1 = \nu + \lambda, \qquad \varkappa_2 = 2(\nu + 2\lambda),$$
$$\varkappa_3 = 8(\nu + 3\lambda), \qquad \varkappa_4 = 48(\nu + 4\lambda).$$

Используя это, доказать, что сумма двух независимых нецентральных χ^2-величин имеет такое же распределение, но с суммарным количеством степеней свободы и суммарным параметром нецентральности.

(Уишарт, 1932; Тэнг, 1938.)

24.2 Показать, что если нецентральные нормальные величины x_i из **24.4** подчиняются k ортогональным линейным ограничениям

$$\sum_{i=1}^{n} a_{ij}x_i = b_j, \qquad j = 1, 2, \ldots, k,$$

где

$$\sum_{i=1}^{n} a_{ij}^2 = 1, \qquad \sum_{i=1}^{n} a_{ij}a_{il} = 0, \qquad j \neq l,$$

то величина

$$y^2 = \sum_{i=1}^{n} x_i^2 - \sum_{j=1}^{k} b_j^2$$

имеет нецентральное χ^2-распределение с $n-k$ степенями свободы и параметром нецентральности $\lambda = \sum_{i=1}^{n} \mu_i^2 - \sum_{j=1}^{k} \left(\sum_{i=1}^{n} a_{ij}\mu_i \right)^2$.

(Патнайк, 1949. См. также Бейтмэн (1949).)

24.3 Доказать, что для любого фиксированного r первые r моментов нецентрального χ^2-распределения с фиксированным λ стремятся при возрастании числа степеней свободы к соответствующим моментам центрального χ^2-распределения с тем же числом степеней свободы. Вывести отсюда, что если статистика критерия имеет нецентральное χ^2-распределение, число степеней свободы которого является возрастающей функцией от объема выборки n, а параметр нецентральности λ равен нулю при гипотезе H_0 и не возрастает, когда $n \to \infty$, при гипотезе H_1, то критерий будет становиться неэффективным при $n \to \infty$, т. е. его мощность будет стремиться к его размеру α.

24.4 Показать, что статистика ОП l, определенная в (24.40) для проверки гипотезы о равенстве k нормальных дисперсий, имеет моменты относительно нуля, равные

$$\mu_r' = \frac{n^{rn/2}\Gamma\{(n-k)/2\}}{\Gamma\{[(r+1)n-k]/2\}} \prod_{i=1}^{k} \frac{\Gamma\{[(r+1)n_i - 1]/2\}}{n_i^{rn_i/2}\Gamma\{(n_i-1)/2\}}.$$

(Нейман и Пирсон, 1931.)

24.5 Пусть проверяется гипотеза H_0 о том, что k нормальных совокупностей имеют одинаковые средние и дисперсии. Показать, что в случае выборок объемов $n_i \geqslant 2$ статистика ОП имеет вид

$$l_0 = \prod_{i=1}^{k} \left(s_i^2/s_0^2\right)^{n_i/2},$$

где

$$s_i^2 = \frac{1}{n_i} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2, \qquad \bar{x} = \frac{1}{n} \sum_{i=1}^{k} n_i \bar{x}_i$$

и

$$s_0^2 = \frac{1}{n} \sum_{i=1}^{k} n_i \left\{ s_i^2 + (\bar{x}_i - \bar{x})^2 \right\}.$$

Показать также, что ее моменты относительно нуля равны

$$\mu_r' = \frac{n^{rn/2}\Gamma\{(n-1)/2\}}{\Gamma\{[(r+1)n-1]/2\}} \prod_{i=1}^{k} \frac{\Gamma\{[(r+1)n_i - 1]/2\}}{n^{rn_i/2}\Gamma\{(n_i-1)/2\}}.$$

(Нейман и Пирсон, 1931.)

24.6 Доказать, что статистика ОП для проверки гипотезы H_2 о равенстве средних в k нормальных совокупностях с одинаковыми дисперсиями имеет вид (при объемах выборок $n_i \geqslant 2$)

$$l_2 = l_0/l,$$

где l и l_0 определены в упражнениях 24.4 и 24.5. Кроме того, доказать, что точное распределение статистики $z = 1 - l_2^{2/n}$, когда выполняется H_2, есть

$$dF \propto z^{(k-3)/2}(1-z)^{(n-k-2)/2}\,dz, \qquad 0 \leqslant z \leqslant 1.$$

Найти моменты величины l_2 и показать, что если выполняется гипотеза H_0 из упражнения 24.5, то l и l_2 независимы.

(Нейман и Пирсон, 1931; Хогг, 1961. См. также Хогг (1962) относительно критерия для гипотезы H_2.)

24.7 Пусть все объемы выборок n_i совпадают. Доказать, что статистика ОП l из (24.40) и ее модифицированная форма l^* из (24.44) связаны соотношением

$$n \log l^* = (n-k) \log l.$$

Таким образом, в этом случае критерии, основанные на l и l^*, эквивалентны.

24.8 Извлекаются выборки из k распределений типа (24.48) или (24.52). Показать, что если l — статистика ОП для проверки гипотезы

$$H_0\colon\ \theta_1 = \theta_2 = \ldots = \theta_{p_1};\quad \theta_{p_1+1} = \theta_{p_1+2} = \ldots = \theta_{p_2};$$
$$\theta_{p_2+1} = \ldots = \theta_{p_3};\ \ldots;\ \theta_{p_{r-1}+1} = \ldots = \theta_{p_r},$$

состоящей в том, что параметры θ_i разбиваются на r различных групп (не обязательно одинакового размера), внутри которых они равны между собой, то величина $-2 \log l$ имеет в точности χ^2-распределение с $2(n-r)$ степенями свободы.

(Хогг, 1956.)

24.9 Пусть имеется выборка объема n из совокупности

$$dF = \frac{dx}{2\theta}, \qquad \mu - \theta \leqslant x \leqslant \mu + \theta.$$

Доказать, что статистика ОП для проверки гипотезы H_0: $\mu = 0$, есть

$$l = \left(\frac{x_{(n)} - x_{(1)}}{2z}\right)^n = \left(\frac{R}{2z}\right)^n,$$

где $z = \max\{-x_{(1)}, x_{(n)}\}$. Используя упражнение 23.7, показать, что l и z независимы, так что имеет место факторизация х. ф.:

$$\mathbf{M}\,[\exp\{(-2\log R^n)\,it\}] = \mathbf{M}\,[\exp\{(-2\log l)\,it\}]\,\mathbf{M}\,[\exp\{[-2\log(2z)^n]\,it\}].$$

Вывести отсюда, что х. ф. величины $-2 \log l$ равна $\varphi = \dfrac{(n-1)}{n(1-2it)-1}$.

Таким образом, когда $n \to \infty$, $-2 \log l$ распределена, как χ^2 с двумя степенями свободы.

(Хогг и Крэйг, 1956.)

24.10 В условиях пункта **24.6** показать, что квадратичная форма $\boldsymbol{x'Ax}$ имеет нецентральное χ^2-распределение тогда и только тогда, когда матрица $\boldsymbol{AV}$ идемпотентна, и что если распределение имеет n степеней свободы, то $\boldsymbol{A} = \boldsymbol{V}^{-1}$.

24.11 Пусть k независимых выборок объемов $n_i \geqslant 2$, $\sum_{i=1}^{k} n_i = n$, извлечены из экспоненциальных распределений

$$dF_i(x) = \exp\left\{-\left(\frac{x-\theta_i}{\sigma_i}\right)\right\}\frac{dx}{\sigma_i}, \qquad \theta_i \leqslant x \leqslant \infty.$$

Показать, что статистика ОП для проверки гипотезы

$$H_0\colon\ \theta_1 = \theta_2 = \ldots = \theta_k, \quad \sigma_1 = \sigma_2 = \ldots = \sigma_k$$

равна

$$l_0 = \prod_{i=1}^{k} d_i^{n_i}/d^n,$$

где $d_i = \bar{x}_i - x_{(1i)}$ есть разность между средним и наименьшим наблюдениями в i-й выборке, а d — такая же статистика для объединенной выборки, т. е. $d = \bar{x} - x_{(1)}$. Показать, что моменты величины $l_0^{1/n}$ равны

$$\mu'_p = \frac{n^p \Gamma(n-1)}{\Gamma(n+p-1)} \prod_{i=1}^{k} \frac{\Gamma\{(n_i-1)+pn_i/n\}}{n_i^{pn_i/n}\Gamma(n_i-1)}.$$

(Сукхатм, 1936.)

24.12 В условиях упражнения 24.11 доказать, что для проверки гипотезы

$$H_1\colon\ \sigma_1 = \sigma_2 = \ldots = \sigma_k,$$

где параметры θ_i не фиксированы, статистика ОП имеет вид

$$l_1 = \frac{\prod_{i=1}^{k} d_i^{n_i}}{\left(\frac{1}{n}\sum_{i=1}^{k} n_i d_i\right)^n}$$

и что моменты функции $l_1^{1/n}$ равны

$$\mu'_p = \frac{n^p \Gamma(n-k)}{\Gamma(n-k+p)} \prod_{i=1}^{k} \frac{\Gamma\{(n_i-1)+pn_i/n\}}{n_i^{pn_i/n}\Gamma(n_i-1)}.$$

(Сукхатм, 1936.)

24.13 Пусть в условиях упражнения 24.11 нам известно, что все σ_i равны. Доказать, что для проверки гипотезы

$$H_2\colon\ \theta_1 = \theta_2 = \ldots = \theta_k$$

статистика ОП имеет вид

$$l_2 = l_0/l_1,$$

где l_0 и l_1 определены в упражнениях 24.11, 24.12. Показать, что точное распределение величины $l_2^{1/n} = u$ есть

$$dF = \frac{1}{\mathrm{B}(n-k,\, k-1)}\, u^{n-k-1}(1-u)^{k-2}\, du, \qquad 0 \leqslant u \leqslant 1,$$

и найти моменты u. Показать, что если выполняется гипотеза H_0 упражнения 24.11, то l_1 и l_2 независимы.

(Сукхатм, 1936; Хогг, 1961. См. также Хогг и Тейнис (1963) относительно других критериев для этих гипотез.)

24.14 Путем сравнения с несмещенным критерием из упражнения 23.17 показать, что критерий ОП для гипотезы о равенстве дисперсий двух нормальных совокупностей является смещенным при неравных объемах выборок n_1, n_2.

24.15 Используя соотношение (24.67), показать, что несмещенный подобный критерий размера α для гипотезы H_0 о равенстве средних у k независимых наблюдений x_i $(i = 1, 2, \ldots, k)$ из нормальных совокупностей с единичными дисперсиями определяется критической областью

$$\sum_{i=1}^{k} (x_i - \bar{x})^2 \geqslant c_\alpha,$$

где c_α есть $100(1-\alpha)$-процентная точка χ^2-распределения с $n-1$ степенями свободы. Показать, что он является также критерием ОП.

24.16 Доказать, что три статистики критериев из упражнений 23.24—23.26 эквивалентны в приведенных ситуациях статистикам ОП, что критическая область критерия ОП в упражнении 23.24 не является РНМН областью и в действительности смещена, но что в двух других упражнениях критерий ОП совпадает с РНМ подобным критерием.

24.17 Обобщая результаты из **23.10—13,** показать, что если распределение имеет вид

$$f(x \mid \theta_1, \theta_2, \ldots, \theta_k) = Q(\boldsymbol{\theta})\, M(x) \exp\left\{ \sum_{j=3}^{k} B_j(x)\, A_j(\theta_3, \theta_4, \ldots, \theta_k) \right\},$$

$$a(\theta_1, \theta_2) \leqslant x \leqslant b(\theta_1, \theta_2)$$

(границы интервала, на котором сосредоточено распределение, зависят только от двух параметров, не входящих в экспоненциальный член), то статистики $t_1 = x_{(1)}$, $t_2 = x_{(n)}$, $t_j = \sum_{i=1}^{n} B_j(x_i)$ являются совместно достаточными для $\boldsymbol{\theta}$ при выборке объема n и их распределение полно.

(Хогг и Крэйг, 1956.)

24.18 Пусть имеются независимые выборки объемов n_1, n_2 из двух распределений

$$dF = \exp\left\{ -\frac{(x_i - \theta_i)}{\sigma} \right\} \frac{dx_i}{\sigma}, \qquad \sigma > 0, \quad x_i \geqslant \theta_i, \quad i = 1, 2.$$

Используя результат упражнения 24.17, показать, что статистики

$$z_1 = \min\{x_{i\,(1)}\}, \qquad z_2 = \sum_{j=1}^{n_1} x_{1j} + \sum_{j=1}^{n_2} x_{2j}$$

являются достаточными для θ_1 и θ_2 и полными.

Показать, что статистика ОП для гипотезы H_0: $\theta_1 = \theta_2$ есть

$$l = \left\{ \frac{z_2 - (n_1 x_{1\,(1)} + n_2 x_{2\,(1)})}{z_2 - (n_1 + n_2) z_1} \right\}^{n_1+n_2}$$

и что l не зависит от z_1, z_2 и, следовательно, от своего знаменателя. Показать, что l дает несмещенный критерий для H_0.

(Полсон, 1941.)

24.19 Обобщая результат упражнения 16.7, показать, что ф. р. нецентрального χ^2-распределения (24.18) дается при четных ν соотношением

$$H(z) = \mathsf{P}\left\{ u - v \geqslant \frac{1}{2}\nu \right\},$$

где u и v — независимые пуассоновские величины с параметрами $z/2$ и $\lambda/2$ соответственно.

(Джонсон, 1959а.)

ГЛАВА 25

СРАВНЕНИЕ КРИТЕРИЕВ

25.1 В главах 22—24 мы занимались отысканием «оптимальных критериев», т. е. критериев, обладающих в данной ситуации «наилучшими» свойствами, где под «наилучшими» понимаются такие желательные свойства, как РНМ, РНМН и др. До сих пор мы не рассматривали вопроса о сравнении, в конкретной ситуации, двух или более критериев с целью оценки их относительной эффективности.

Исследование этого вопроса необходимо для оценивания потери эффективности, возникающей при использовании критерия, отличного от оптимального. Может случиться, например, что РНМ критерий имеет незначительно бо́льшую мощность, чем какой-либо другой критерий, который, возможно, много проще с точки зрения вычисления. В такой ситуации вполне естественно для стандартной проверки гипотезы использовать менее эффективный критерий. Но прежде чем принять такое решение, мы должны сделать какое-то количественное сравнение между критериями.

Аналогичная проблема обсуждалась в **17.29** в теории оценивания, где была получена мера эффективности оценки. Читатель, возможно, спросит, как случилось, что в теории оценивания вопрос измерения эффективности обсуждался сразу же по введении понятия эффективности, а здесь вопрос измерения эффективности критерия отнесен в конец общего обсуждения теории критериев. Ответ частично состоит в том, что понятие эффективности критерия оказывается значительно более сложным, чем понятие эффективности оценки, и требует поэтому более развернутого изложения.

Основная причина, однако, заключается в том, что мы попросту следуем историческому развитию этого вопроса: необходимость измерения эффективности критериев возникла только тогда, когда (начиная примерно с 1935 г.) внимание статистиков было обращено на простые с точки зрения вычислений критерии (они будут обсуждаться в главах 31 и 32) Даже идея состоятельности критерия, с которой мы встретились в **24.17**, была развита Вальдом и Волфовицем (1940) только почти через двадцать

лет после первого определения Фишером (1921а) состоятельной оценки. Лишь когда «неэффективные» критерии стали представлять практический интерес, возникла необходимость исследования более слабых свойств критериев.

Сравнение функций мощности

25.2 При проверке данной гипотезы против данной альтернативы, когда объем выборки фиксирован, простейший путь сравнения двух критериев состоит в непосредственном исследовании их функции мощности. Если объем выборки находится в нашем распоряжении (например, мы планируем серию наблюдений), то естественно искать определение эффективности критерия в такой же форме, как для эффективности оценки в **17.19**. Пусть некоторому «эффективному» критерию (т. е. наиболее мощному в рассматриваемом классе) размера α необходимо n_1 наблюдений для достижения определенной мощности против данной альтернативы, а другому критерию того же размера α требуется для этого n_2 наблюдений. Тогда можно определить *относительную эффективность* второго критерия для достижения указанной мощности против заданной альтернативы как n_1/n_2. Введенная мера представляет собой, как и в случае оценивания, величину, обратную к отношению требуемых объемов выборок. Заметим, что наше определение относительной эффективности не является асимптотическим и не содержит никаких ограничений на выборочные распределения сравниваемых статистик критериев. Таким путем можно сравнивать любые два критерия, потому что функции мощности критериев, для которых вычисляется относительная эффективность, полностью учитывают распределения статистик критериев (функции мощности содержат в себе всю информацию, относящуюся к сравнению).

Асимптотические сравнения

25.3 Понятие относительной эффективности дает исчерпывающую информацию, но оказывается неудобным для применения. Подобно функции мощности, на которой оно основано, это понятие представляет собой функцию трех аргументов — размера критерия α, «расстояния» (в терминах некоторого параметра θ) между проверяемой гипотезой и альтернативой, а также объема выборки n_1, требуемого для эффективного критерия. Если даже ограничиться несколькими типичными значениями α, то для сравнения критериев с помощью этой меры требуется таблица с двумя входами. Было бы значительно удобнее, если бы мы смогли найти некую единую результирующую меру эффективности. Понятно, что можно надеяться достичь этого, лишь имея

дело с некоторым предельным процессом. Таким образом, мы вернулись к необходимости ограничиться асимптотическими результатами.

25.4 Другой подход, который напрашивается сам собой, состоит в том, чтобы устремить объем выборки к бесконечности, как в **17.29**, и взять отношение мощностей критериев в качестве меры эффективности критерия. Если мы попытаемся воспользоваться этим подходом, то немедленно столкнемся с трудностью. Действительно, когда рассматриваются два состоятельных против данного класса альтернативных гипотез критерия размера α (а мы всегда будем иметь дело только с таким случаем), то функция мощности каждого из них по определению стремится к 1 с возрастанием объема выборки. Отсюда следует, что если мы сравниваем критерий против некоторого фиксированного альтернативного значения параметра θ, то определенная таким образом эффективность всегда будет стремиться к 1 при возрастании объема выборки. Приведенная мера эффективности оказывается, следовательно, совершенно бесполезной.

Вообще, легко видеть, что асимптотическое по n рассмотрение функций мощности состоятельных критериев имеет ограниченное значение. Например, Вальд (1941) определил асимптотически наиболее мощный критерий как критерий, у которого функция мощности не может быть улучшена, когда объем выборки стремится к бесконечности, т. е. который является асимптотически РНМ критерием. Следующий пример, принадлежащий Леману (1949), показывает, что один асимптотически РНМ критерий может быть (даже асимптотически) явно хуже, чем другой такой же критерий.

Пример 25.1

Рассмотрим вновь задачу проверки гипотезы о среднем θ нормальной совокупности с известной дисперсией, равной 1, обсуждавшуюся в примерах 22.1 и 22.2. Мы хотим проверить гипотезу H_0: $\theta = \theta_0$ против односторонней альтернативы H_1: $\theta = \theta_1 > \theta_0$. В **22.17** мы видели, что РНМ критерий для H_0 против H_1 задается критической областью $\bar{x} \geqslant \theta_0 + d_\alpha/n^{1/2}$, а в примере 22.3 было показано, что его функция мощности равна

$$P_1 = G\{\Delta n^{1/2} - d_\alpha\} = 1 - G\{d_\alpha - \Delta n^{1/2}\}, \qquad (25.1)$$

где $\Delta = \theta_1 - \theta_0$ и фиксированное значение d_α определяет, как в (22.16), размер критерия α.

Построим теперь двусторонний критерий уровня α, отклоняющий гипотезу H_0, когда

$$\bar{x} \geqslant \theta_0 + d_{\alpha_2}/n^{1/2} \quad \text{или} \quad \bar{x} \leqslant \theta_0 - d_{\alpha_1}/n^{1/2},$$

где значения d_{α_1} и d_{α_2} (зависящие от n) могут быть выбраны произвольно, лишь бы только $\alpha_1 + \alpha_2 = \alpha$, откуда следует, что d_{α_1} и d_{α_2} превосходят d_α. Согласно (23.48) функция мощности второго критерия равна

$$P_2 = G\{\Delta n^{1/2} - d_{\alpha_2}\} + G\{-\Delta n^{1/2} - d_{\alpha_1}\}, \tag{25.2}$$

а из положительности G вытекает, что

$$P_2 > G\{\Delta n^{1/2} - d_{\alpha_2}\} = 1 - G\{d_{\alpha_2} - \Delta n^{1/2}\}. \tag{25.3}$$

Так как первый критерий является РНМ, то, вычитая (25.3) из (25.1), получаем

$$G\{d_{\alpha_2} - \Delta n^{1/2}\} - G\{d_\alpha - \Delta n^{1/2}\} > P_1 - P_2 \geqslant 0. \tag{25.4}$$

Легко видеть, что разность между $G(x)$ и $G(y)$ при фиксированной величине $x - y$ будет максимальной, когда x и y симметричны относительно нуля, т. е. когда $x = -y$. Отсюда следует, что

$$G\left\{\frac{1}{2}(x-y)\right\} - G\left\{-\frac{1}{2}(x-y)\right\} \geqslant G\{x\} - G\{y\}. \tag{25.5}$$

Учитывая (25.5), имеем из (25.4)

$$G\left\{\frac{1}{2}(d_{\alpha_2} - d_\alpha)\right\} - G\left\{-\frac{1}{2}(d_{\alpha_2} - d_\alpha)\right\} > P_1 - P_2 \geqslant 0. \tag{25.6}$$

Таким образом, если для каждого n выбрать d_{α_2} так, чтобы

$$\lim_{n\to\infty} d_{\alpha_2} = d_\alpha, \tag{25.7}$$

то левая часть (25.6) будет стремиться к нулю, и, значит $P_1 - P_2$ будет стремиться к нулю равномерно по Δ. Следовательно, такой двусторонний критерий является также асимптотически РНМ критерием.

Рассмотрим теперь для этих критериев отношение ошибок второго рода. Из (25.1) и (25.2) имеем

$$\frac{1-P_2}{1-P_1} = \frac{G\{d_{\alpha_2} - \Delta n^{1/2}\} - G\{-d_{\alpha_1} - \Delta n^{1/2}\}}{G\{d_\alpha - \Delta n^{1/2}\}}. \tag{25.8}$$

Когда $n^{1/2} \to \infty$, числитель и знаменатель в (25.8) стремятся к нулю. По правилу Лопиталя (обозначая штрихом дифференцирование по $n^{1/2}$ и буквой g нормальную ф. п.) находим

$$\lim_{n^{1/2}\to\infty} \frac{1-P_2}{1-P_1} =$$

$$= \lim_{n^{1/2}\to\infty} \left[\frac{(d'_{\alpha_2} - \Delta)\, g\{d_{\alpha_2} - \Delta n^{1/2}\}}{-\Delta g\{d_\alpha - \Delta n^{1/2}\}} + \frac{(d'_{\alpha_1} + \Delta)\, g\{-d_{\alpha_1} - \Delta n^{1/2}\}}{-\Delta g\{d_\alpha - \Delta n^{1/2}\}} \right]. \tag{25.9}$$

Из (25.7) следует, что $d_{a_1} \to \infty$ вместе с n, и поэтому второе слагаемое в правой части (25.9) стремится к нулю. Кроме того, из (25.7) следует, что первое слагаемое в правой части (25.9) будет стремиться к бесконечности, если предел

$$\lim_{n\to\infty} \frac{-d'_{a_2} g\{d_{a_2} - \Delta n^{1/2}\}}{g\{d_a - \Delta n^{1/2}\}} = \lim_{n\to\infty} \left[-d'_{a_2} \frac{\exp\left\{-\frac{1}{2}(d_{a_2} - \Delta n^{1/2})^2\right\}}{\exp\left\{-\frac{1}{2}(d_a - \Delta n^{1/2})^2\right\}} \right] =$$

$$= \lim_{n\to\infty} \left[-d'_{a_2} \exp\left\{ -\frac{1}{2}(d_{a_2}^2 - d_a^2) + \Delta n^{1/2}(d_{a_2} - d_a) \right\} \right] \quad (25.10)$$

равен бесконечности. Согласно (25.7) первое слагаемое в экспоненте в правой части (25.10) стремится к нулю. Если положить

$$d_{a_2} = d_a + n^{-\delta}, \qquad 0 < \delta < \frac{1}{2}, \qquad (25.11)$$

то соотношение (25.7) будет выполняться, а (25.10) будет иметь своим пределом по n бесконечность. Таким образом, отношение ошибок второго рода (25.8) стремится к бесконечности вместе с n, хотя оба критерия являются асимптотически РНМ. Отсюда видно, что это асимптотическое свойство критериев оказывается не очень избирательным.

Асимптотическая относительная эффективность

25.5 Для того чтобы из относительной эффективности получить полезную асимптотическую меру эффективности критерия, рассмотрим предел относительных эффективностей критериев, в которых с ростом n альтернативное значение параметра θ приближается к проверяемому значению θ_0. Этот тип альтернатив был впервые исследован Питмэном (1948), результаты которого были обобщены Нётером (1955). Другие типы предельных процессов в определении относительных эффективностей рассматривались Диксоном (1953b), Ходжесом и Леманом (1956).

Пусть t_1 и t_2 — статистики состоятельных критериев для проверки гипотезы H_0: $\theta = \theta_0$ против односторонней альтернативы H_1: $\theta > \theta_0$. Предположим на время, что t_1 и t_0 асимптотически нормальны при любом значении θ, — в **25.14—15** мы ослабим это требование. Для краткости обозначим

$$\mathsf{M}(t_i \mid H_j) = E_{ij}, \quad \mathsf{D}(t_i \mid H_j) = D_{ij}^2, \quad E_i^{(r)}(\theta) = \frac{\partial^r}{\partial\theta^r} E_{i1},$$

$$D_{i1}^{(r)} = \frac{\partial^r}{\partial\theta^r} D_{i1}, \quad D_{i0}^{(r)} = D_{i1}^{(r)}(\theta_0) \qquad (i = 1,\ 2;\quad j = 0,\ 1).$$

При больших выборках критерии размера α определяются критическими областями

$$t_i > E_{i0} + \lambda_\alpha D_{i0} \tag{25.12}$$

(знак у t_i выбирается так, чтобы область имела такой вид), где λ_α — нормальное отклонение, определяемое из уравнения $G\{-\lambda_\alpha\} = \alpha$, а G, как и прежде, — стандартная нормальная ф. р. Следуя в точности примеру 22.3, для асимптотической функции мощности, основанной на статистике t_i, имеем

$$P_i(\theta) = G\{[E_{i1} - (E_{i0} + \lambda_\alpha D_{i0})]/D_{i1}\}. \tag{25.13}$$

Обозначая $u_i(\theta, \lambda_\alpha)$ аргумент функции G в (25.13) и разлагая $E_{i1} - E_{i0}$ в ряд Тейлора, получаем

$$u_i(\theta, \lambda_\alpha) = \left[E_i^{(m_i)}(\theta_i^*)\frac{(\theta - \theta_0)^{m_i}}{m_i!} - \lambda_\alpha D_{i0}\right]\Big/ D_{i1}, \tag{25.14}$$

где $\theta_0 < \theta_i^* < \theta$, а m_i — порядок первой ненулевой производной в точке θ_0, т. е. m_i определяется условиями

$$\left.\begin{aligned} &E_i^{(r)}(\theta_0) = 0, \quad r = 1, 2, \ldots, m_i - 1, \\ &E_i^{(m_i)}(\theta_0) \neq 0. \end{aligned}\right\} \tag{25.15}$$

С целью задания альтернативной гипотезы предположим, что при $n \to \infty$

$$R_i = \left[E_i^{(m_i)}(\theta_0)/D_{i0}\right] \sim c_i n^{m_i \delta_i}. \tag{25.16}$$

Соотношение (25.16) определяет константы $\delta_i > 0$ и c_i. Рассмотрим теперь последовательности альтернативных значений параметра, приближающихся к θ_0 при $n \to \infty$,

$$\theta = \theta_0 + k_i/n^{\delta_i}, \tag{25.17}$$

где k_i — произвольная положительная константа. Если выполнены условия регулярности

$$\lim_{n\to\infty} \frac{E_i^{(m_i)}(\theta)}{E_i^{(m_i)}(\theta_0)} = 1, \qquad \lim_{n\to\infty} \frac{D_{i1}}{D_{i0}} = 1, \tag{25.18}$$

то с помощью (25.16), (25.17) и (25.18) равенство (25.14) приводится к виду

$$u_i(\theta, \lambda_\alpha) = \frac{c_i k_i^{m_i}}{m_i!} - \lambda_\alpha, \tag{25.19}$$

и согласно (25.13) асимптотические мощности критериев равны $G\{u_i\}$.

25.6 Для того чтобы при любом фиксированном α два критерия имели одинаковые мощности против одной и той же последовательности альтернатив, согласно (25.17) и (25.19) должно быть

$$\frac{k_1}{n_1^{\delta_1}} = \frac{k_2}{n_2^{\delta_2}} \tag{25.20}$$

и

$$\frac{c_1 k_1^{m_1}}{m_1!} = \frac{c_2 k_2^{m_2}}{m_2!}, \tag{25.21}$$

где n_1 и n_2 — объемы выборок, на которых основаны t_1 и t_2. Объединяя (25.20) и (25.21), имеем

$$\frac{n_1^{\delta_1}}{n_2^{\delta_2}} = \left(\frac{c_2}{c_1} \frac{m_1!}{m_2!} k_2^{m_2 - m_1} \right)^{1/m_1}. \tag{25.22}$$

Правая часть в (25.22) является положительной константой. Таким образом, если n_1, $n_2 \to \infty$, то отношение n_1/n_2 будет стремиться к константе тогда и только тогда, когда $\delta_1 = \delta_2$. Если $\delta_1 > \delta_2$, то мы должны иметь $n_1/n_2 \to 0$, в случае же $\delta_1 < \delta_2$ имеем $n_1/n_2 \to \infty$. Определяя *асимптотическую относительную эффективность* (АОЭ) критерия t_2 по сравнению с t_1 как

$$A_{21} = \lim \frac{n_1}{n_2}, \tag{25.23}$$

мы, следовательно, получаем, что

$$A_{21} = 0, \quad \text{если} \quad \delta_1 > \delta_2. \tag{25.24}$$

Таким образом, для сравнения двух критериев с помощью АОЭ нужно вначале сравнить соответствующие им значения δ: если у одного из них δ меньше, чем у другого, то первый критерий имеет нулевую АОЭ по сравнению со вторым. Величина δ играет здесь такую же роль, что и порядок малости дисперсии при измерении эффективности оценки (см. **17.29**).

Мы можем теперь ограничиться случаем $\delta_1 = \delta_2 = \delta$. Тогда соотношения (25.22) и (25.23) дают

$$A_{21} = \lim \frac{n_1}{n_2} = \left(\frac{c_2}{c_1} \frac{m_1!}{m_2!} k^{m_2 - m_1} \right)^{1/(m_1 \delta)}. \tag{25.25}$$

Если вдобавок

$$m_1 = m_2 = m, \tag{25.26}$$

то (25.25) приводится к виду

$$A_{21} = \left(\frac{c_2}{c_1} \right)^{1/(m\delta)}.$$

Из последнего соотношения, учитывая (25.16), находим

$$A_{21} = \lim_{n\to\infty} \left\{ \frac{E_2^{(m)}(\theta_0)/D_{20}}{E_1^{(m)}(\theta_0)/D_{10}} \right\}^{1/(m\delta)}. \qquad (25.27)$$

В большинстве задач величину (25.27) легко оценить, и мы часто будем ее использовать в последующих главах для получения АОЭ конкретных критериев. Чаще всего $\delta = 1/2$ (что соответствует оценке дисперсии порядка n^{-1}) и $m = 1$. Интерпретацию величины m см. ниже, в **25.10.**

Можно заметить, что если $m_2 \neq m_1$, то выражение (25.25) неопределенно и зависит от произвольной постоянной k_2. Поэтому критерии с равными значениями δ имеют разные АОЭ при различных последовательностях альтернатив (25.17), если только значения m не одинаковы. Причины этого будут указаны в **25.10.**

25.7 Если мы проверяем гипотезу H_0 против двусторонней альтернативы H_1: $\theta \neq \theta_0$, то наши результаты об АОЭ остаются в силе, если использовать критические области с «равными хвостами» вида

$$t_i > E_{i0} + \lambda_{\alpha/2} D_{i0} \quad \text{или} \quad t_i < E_{i0} - \lambda_{\alpha/2} D_{i0}.$$

Действительно, в этом случае асимптотические функции мощности (25.13) заменяются на

$$Q_i(\theta) = G\{u_i(\theta, \lambda_{\alpha/2})\} + 1 - G\{u_i(\theta, -\lambda_{\alpha/2})\}, \qquad (25.28)$$

и, как и прежде, $Q_1 = Q_2$ против альтернативы (25.17) (где k_i уже могут быть любого знака), если выполняются (25.20) и (25.21). Конейн (1956) приводит более общее исследование двусторонних критериев, которые не обязательно имеют «равные хвосты».

Пример 25.2

Пусть проверяется гипотеза о среднем θ нормальной совокупности с известной дисперсией σ^2. Сравним критерий, основанный на выборочной медиане $\tilde{x}$, с РНМ критерием, основанным на выборочном среднем $\bar{x}$. Обе статистики асимптотически нормальны. Известно, что

$$\mathsf{M}(\bar{x}) = \theta, \qquad \mathsf{D}(\bar{x} \mid \theta) = \frac{\sigma^2}{n},$$

а $\tilde{x}$ — состоятельная оценка параметра θ, у которой (см. в примере 10.7)

$$\mathsf{M}(\tilde{x}) = \theta, \quad \mathsf{D}(\tilde{x} \mid \theta) \sim \pi\sigma^2/(2n).$$

Таким образом, для обоих критериев имеем

$$E'(\theta_0) = 1,$$

откуда $m_1 = m_2 = 1$, в то же время согласно (25.16) $\delta_1 = \delta_2 = 1/2$. Следовательно, из (25.27) находим

$$A_{\tilde{x},\,\bar{x}} = \lim_{n\to\infty} \left\{ \frac{1/(\pi\sigma^2/2n)^{1/2}}{1/(\sigma^2/n)^{1/2}} \right\}^2 = \frac{2}{\pi}.$$

Тот же самый результат был получен в примере 17.12 для эффективности медианы $\tilde{x}$ при оценке параметра θ. В **25.13** мы увидим, что это частный случай общей взаимосвязи между эффективностью оценки и АОЭ критериев.

АОЭ и производные функций мощности

25.8 Природа последовательности альтернативных гипотез (25.7), приближающейся к θ_0 при $n \to \infty$, показывает, что АОЭ каким-то образом связана с поведением функций мощности сравниваемых критериев вблизи θ_0. Мы установим эту связь точнее, показав, что при определенных условиях АОЭ просто выражается через отношение производных функций мощностей.

Рассмотрим сначала случай односторонней гипотезы H_1, обсуждавшийся в **25.5—6**. В этом случае функции мощности критериев задаются асимптотически соотношением (25.13), которое, как и прежде, запишем в виде

$$P_i(\theta) = G\{u_i(\theta, \lambda_\alpha)\}. \tag{25.29}$$

Дифференцируя по θ, имеем

$$P_i'(\theta) = g\{u_i\}\, u_i'(\theta, \lambda_\alpha), \tag{25.30}$$

где g — нормальная функция плотности. Из (25.13) находим

$$u_i'(\theta, \lambda_\alpha) = \frac{E_{i1}'}{D_{i1}} - \frac{D_{i1}'}{D_{i1}^2}(E_{i1} - E_{i0} - \lambda_\alpha D_{i0}). \tag{25.31}$$

Используя (25.18) и дополнительные условия регулярности

$$\lim_{n\to\infty} \frac{D_{i1}'}{D_{i0}'} = 1, \quad \lim_{n\to\infty} \frac{E_{i1}}{E_{i0}} = 1,$$

получаем, что (25.31) при $n \to \infty$ превращается в

$$u_i'(\theta, \lambda_\alpha) = \frac{E_i'(\theta_0)}{D_{i0}} + \frac{D_{i0}'}{D_{i0}}\lambda_\alpha. \tag{25.32}$$

Таким образом, если в (25.15) $m_i = 1$ и если

$$\lim_{n\to\infty} \frac{D'_{i0}}{E'_i(\theta_0)} = 0, \tag{25.33}$$

то (25.32) в точке θ_0 приводится к виду

$$u'_i(\theta_0, \lambda_\alpha) \sim E'_i(\theta_0)/D_{i0}. \tag{25.34}$$

Поскольку согласно (25.13)

$$g\{u_i(\theta_0, \lambda_\alpha)\} = g\{-\lambda_\alpha\}, \tag{25.35}$$

то, подставляя (25.34) и (25.35) в (25.30), имеем

$$P'_i(\theta_0) = P'_i(\theta_0, \lambda_\alpha) \sim g\{-\lambda_\alpha\} E'_i(\theta_0)/D_{i0}. \tag{25.36}$$

Вспоминая, что $m_i = 1$, из (25.36) и (25.27) находим

$$\lim_{n\to\infty} \frac{P'_2(\theta_0)}{P'_1(\theta_0)} = A_{21}^{\delta}, \qquad m_1 = m_2 = 1, \tag{25.37}$$

т. е. асимптотическое отношение первых производных функций мощности критериев в точке θ_0 равно АОЭ в степени δ (обычно $\delta = 1/2$). Таким образом, если бы мы использовали это отношение как показатель асимптотической эффективности критериев, то получали бы в точности такие же результаты, как при использовании АОЭ. Этот показатель был предложен (под названием «асимптотическая локальная эффективность») Бломквистом (1950).

25.9 Если $m_i > 1$, т. е. $E'_i(\theta_0) = 0$, то (25.36) равно нулю (при нашей степени аппроксимации) и результат из **25.8** бесполезен. Нужно продолжить дифференцирование для того, чтобы получить полезные результаты.

Из (25.30) находим

$$P''_i(\theta) = \frac{\partial g\{u_i\}}{\partial u_i}[u'_i(\theta, \lambda_\alpha)]^2 + g\{u_i\} u''_i(\theta, \lambda_\alpha). \tag{25.38}$$

Согласно (25.31)

$$u''_i(\theta, \lambda_\alpha) =$$

$$= \frac{E''_{i1}}{D_{i1}} - \frac{2E'_{i1}D'_{i1}}{D_{i1}^2} - (E_{i1} - E_{i0} - \lambda_\alpha D_{i0})\left[\frac{D''_{i1}}{D_{i1}^2} - \frac{2(D'_{i1})^2}{D_{i1}^3}\right]. \tag{25.39}$$

Если выполняются условия (25.18) с $m_i = 2$, условия регулярности, указанные после (25.31), а также соотношения

$$\lim_{n\to\infty} E'_{i1} = E'_i(\theta_0) = 0, \quad \lim \frac{D'_{i1}}{D'_{i0}} = 1, \quad \lim_{n\to\infty} \frac{D''_{i1}}{D''_{i0}} = 1, \quad \lim_{n\to\infty} \frac{E_{i1}}{E_{i0}} = 1, \tag{25.40}$$

то равенство (25.39) принимает вид

$$u_i''(\theta_0, \lambda_\alpha) = \frac{E_i''(\theta_0)}{D_{i0}} + \lambda_\alpha \left[\frac{D_{i0}''}{D_{i0}} - 2\left(\frac{D_{i0}'}{D_{i0}}\right)^2 \right]. \qquad (25.41)$$

Вместо (25.33) предположим теперь, что

$$\lim_{n\to\infty} \frac{D_{i0}''}{E_i''(\theta_0)} = 0, \quad \lim_{n\to\infty} \frac{(D_{i0}')^2}{D_{i0}E_i''(\theta_0)} = 0. \qquad (25.42)$$

Условия (25.42) приводят (25.41) к виду

$$u_i''(\theta_0, \lambda_\alpha) \sim E_i''(\theta_0)/D_{i0}. \qquad (25.43)$$

Возвращаясь к (25.38), мы видим, что поскольку

$$\frac{\partial g\{u_i\}}{\partial u_i} = -u_i g\{u_i\},$$

то, используя (25.32), (25.35) и (25.43), из (25.38) имеем

$$P_i''(\theta_0) \sim g\{-\lambda_\alpha\} \left\{ \lambda_\alpha \left[\frac{E_i'(\theta_0)}{D_{i0}} + \frac{D_{i0}'}{D_{i0}} \lambda_\alpha \right]^2 + \frac{E_i''(\theta_0)}{D_{i0}} \right\}. \qquad (25.44)$$

Так как мы рассматриваем случай $m_i = 2$, то член с $E_i'(\theta_0)$ равен нулю. Учитывая второе из условий (25.42), перепишем окончательно соотношение (25.44) в виде

$$P_i''(\theta_0) \sim g\{-\lambda_\alpha\} E_i''(\theta_0)/D_{i0}. \qquad (25.45)$$

Отсюда при $m = 2$ для предела отношения вторых производных равенство (25.27) дает

$$\lim_{n\to\infty} \frac{P_2''(\theta_0)}{P_1''(\theta_0)} = A_{21}^{2\delta}. \qquad (25.46)$$

Соотношения (25.37) и (25.46) можно объединить в следующем утверждении. Для $m = 1, 2$ отношение m-х производных функций мощности односторонних критериев асимптотически равно АОЭ, возведенной в степень $m\delta$.

Если вместо (25.33) и (25.42) наложить более сильные условия

$$\lim_{n\to\infty} D_{i0}'/D_{i0} = 0, \quad \lim_{n\to\infty} D_{i0}''/D_{i0} = 0, \qquad (25.47)$$

которые вместе с (25.16) влекут (25.33) и (25.42), то (25.34) и (25.43) будут следовать, как и прежде, из (25.32) и (25.41). В конкретных случаях предположения (25.47) могут оказаться более простыми для проверки.

Интерпретация величины m

25.10 Обсудим теперь общие условия, при которых m принимает значение 1 или 2. Рассмотрим снова асимптотическую функцию мощности (25.13) критерия (25.12) для односторонней альтернативы H_1: $\theta_1 > \theta_0$. Для краткости в этом пункте мы будем опускать индекс i. Если $\theta \to \theta_0$ и по условию (25.18) $D_1 \to D_0$, то

$$P(\theta) = G\left\{\frac{E_1 - E_0}{D_0} - \lambda_\alpha\right\}$$

есть монотонно возрастающая функция от $(E_1 - E_0)$.

Если $(E_1 - E_0)$ является возрастающей функцией от $(\theta - \theta_0)$, то $P(\theta) \to 0$, когда $\theta \to -\infty$ (последнее означает, что в качестве критической области при $\theta < \theta_0$ использовался бы другой «хвост» распределения статистики критерия). Если существует производная $E'(\theta_0)$, то она отлична от нуля и $m = 1$, а также согласно (25.36) $P'(\theta_0) \neq 0$.

Предположим, с другой стороны, что $(E_1 - E_0)$ представляет собой четную функцию от $(\theta - \theta_0)$, возрастающую при росте $|\theta - \theta_0|$ (в этом случае в качестве критической области при любом знаке разности $(\theta - \theta_0)$ использовался бы один и тот же «хвост»). Пусть, кроме того, существует производная $E'(\theta_0)$, которая должна, при условиях регулярности, равняться нулю. Тогда $m > 1$ и в практических случаях оказывается, что $m = 2$. Согласно (25.36) также $P'(\theta_0) = 0$ с той же степенью аппроксимации.

Мы теперь можем понять, почему (как замечено в конце **25.6**) АОЭ бесполезна при сравнении критериев с различными значениями m, равными на практике 1 и 2. Это происходит потому, что тогда сравниваются критерии, функции мощности которых ведут себя существенно по-разному в точке θ_0: одна имеет там регулярный минимум, другая — нет. Неопределенность выражения (25.25) не является для такой ситуации неожиданной. Нужно добавить, что в данный момент эта неопределенность представляет чисто теоретический интерес, поскольку, кажется, неизвестны случаи, к которым она была бы приложима.

Пример 25.3

Рассмотрим задачу проверки гипотезы H_0: $\theta = \theta_0$ в случае нормальной совокупности со средним θ и дисперсией 1. Мы имеем пару односторонних РНМ критериев, основанных на выборочном среднем $\bar{x}$ (см. **22.17**), причем в качестве критической области выбирается верхний или нижний «хвост», когда H_1 со-

стоит из $\theta > \theta_0$ или $\theta < \theta_0$ соответственно. Согласно примеру 25.2, для $\bar{x}$ $\delta = 1/2$ и $m = 1$.

В качестве статистики критерия можно было бы использовать также

$$S = \sum_{i=1}^{n} (x_i - \theta_0)^2.$$

Величина S имеет нецентральное χ^2-распределение с n степенями свободы и параметром нецентральности $n(\theta - \theta_0)^2$; поэтому (см. упражнение 24.1)

$$\mathsf{M}(S \mid \theta) = n\{1 + (\theta - \theta_0)^2\}, \qquad \mathsf{D}(S \mid \theta_0) = 2n.$$

Когда $n \to \infty$, статистика S асимптотически нормальна. Имеем $E'(\theta) = 2n(\theta - \theta_0)$, $E'(\theta_0) = 0$, $E''(\theta) = 2n = E''(\theta_0)$, откуда $m = 2$ и

$$\frac{E''(\theta_0)}{D_0} = \frac{2n}{(2n)^{1/2}} = (2n)^{1/2}.$$

Поскольку $m = 2$, то из (25.16) $\delta = 1/4$. Так как для $\bar{x}$ $\delta = 1/2$, то согласно (25.24) АОЭ статистики S по сравнению с $\bar{x}$ равна нулю. Критическая область для S образована верхним «хвостом», независимо от значения θ.

25.11 Перейдем теперь к случаю двусторонней альтернативы H_1: $\theta \neq \theta_0$. Функция мощности критерия с «равными хвостами» дается асимптотически соотношением (25.28). Ее производная в точке θ_0 равна

$$Q_i'(\theta_0) = P_i'(\theta_0, \lambda_{\alpha/2}) - P_i'(\theta_0, -\lambda_{\alpha/2}), \tag{25.48}$$

где P_i' определяются из (25.36), если $m_i = 1$ и выполняется (25.33) или (25.47). Поскольку в (25.36) $g\{-\lambda_\alpha\}$ является четной функцией относительно λ_α, то из (25.48) немедленно получаем асимптотический результат

$$Q_i'(\theta_0) \sim 0,$$

говорящий о том, что наклон функции мощности в θ_0 асимптотически равен нулю. Этот результат следует также (при условиях регулярности) из замечания в **24.17**, касающегося асимптотической несмещенности состоятельных критериев.

Вторая производная функции мощности (25.28) равна

$$Q_i''(\theta_0) = P_i''(\theta_0, \lambda_{\alpha/2}) - P_i''(\theta_0, -\lambda_{\alpha/2}). \tag{25.49}$$

Мы нашли P_i'' в (25.44), где было $m_i = 2$. Соотношение (25.44) будет выполняться и при $m_i = 1$, если усилить первое условие в (25.47), полагая

$$D_{i0}'/D_{i0} = o(n^{-\delta}), \tag{25.50}$$

потому что тогда, согласно (25.16), вторым членом в правой

части (25.39) можно пренебречь и мы получаем, как и прежде, (25.44). Подставляя (25.44) в (25.49), имеем

$$Q_i''(\theta_0) \sim 2\lambda_{\alpha/2} g\{-\lambda_{\alpha/2}\}\left\{\left(\frac{E_i'(\theta_0)}{D_{i0}}\right)^2 + \left(\frac{D_{i0}'}{D_{i0}}\lambda_\alpha\right)^2\right\},$$

а с помощью условия (25.50) это соотношение преобразуется к виду

$$Q_i''(\theta_0) \sim 2\lambda_{\alpha/2} g\{-\lambda_{\alpha/2}\}\left(\frac{E_i'(\theta_0)}{D_{i0}}\right)^2. \tag{25.51}$$

Следовательно, в этом случае из (25.27) и (25.51) находим

$$\frac{Q_2''(\theta_0)}{Q_1''(\theta_0)} = A_{21}^{2\delta}. \tag{25.52}$$

Таким образом, если $m = 1$, то асимптотическое отношение вторых производных функций мощности двусторонних критериев совпадает с соответствующим отношением (25.46) для односторонних критериев, когда $m = 2$. Оно равно также квадрату выражения (25.37), относящегося к односторонним критериям при $m = 1$.

Случай $m = 2$ для критериев с двусторонней критической областью кажется не очень важным: замечания в **25.10** показывают, что там, где $m = 2$, часто может быть использован критерий с односторонней критической областью даже против двусторонней альтернативы H_1.

Пример 25.4

В примере 25.2 оба критерия имеют $\delta = 1/2$, $m = 1$ и $E'(\theta_0) = 1$. Поскольку дисперсия каждой статистики не зависит от θ, по крайней мере асимптотически, то выполнены предположения (25.33) и (25.50) и условия регулярности, так что из (25.37) для односторонних критериев имеем

$$\lim_{n\to\infty} \frac{P_{\bar{x}}'(\theta_0)}{P_{\tilde{x}}'(\theta_0)} = A_{\tilde{x},\bar{x}}^{1/2} = \left(\frac{2}{\pi}\right)^{1/2},$$

тогда как для двусторонних критериев из (25.52) получаем

$$\lim_{n\to\infty} \frac{Q_{\bar{x}}''(\theta_0)}{Q_{\tilde{x}}''(\theta_0)} = A_{\bar{x},\tilde{x}} = \frac{2}{\pi}.$$

АОЭ и максимальная потеря мощности

25.12 Хотя АОЭ критериев отражает в основном свойства их мощности в окрестности точки θ_0, она имеет также некоторое отношение к асимптотической функции мощности в целом, по крайней мере в случае $m = 1$, которым мы теперь ограни-

чимся. Функция мощности $P_i(\theta)$ одностороннего критерия равна $G\{u_i(\theta)\}$, где функция $u_i(\theta)$ определена в (25.14) и при условиях регулярности (25.18), когда $m_i = 1$, асимптотически равна

$$u_i(\theta) = \frac{E_i'(\theta_0)}{D_{i0}}(\theta - \theta_0) - \lambda_\alpha. \qquad (25.53)$$

Таким образом, $u_i(\theta)$ асимптотически линейна по θ. Если, как в (25.16), обозначим $R_i = E_i'(\theta_0)/D_{i0}$, то разность между двумя такими функциями мощности можно представить в виде

$$d(\theta) = P_2(\theta) - P_1(\theta) = \\ = G\{(\theta - \theta_0) R_2 - \lambda_\alpha\} - G\left\{(\theta - \theta_0) R_2 \frac{R_1}{R_2} - \lambda_\alpha\right\}, \qquad (25.54)$$

где без потери общности предполагается, что $R_2 > R_1$. Рассмотрим поведение $d(\theta)$ как функции от θ. Когда $\theta = \theta_0$, $d = 0$. При θ, стремящемся к бесконечности, P_1 и P_2 стремятся к 1, а d — вновь к нулю. Максимальное значение функции $d(\theta)$ зависит только от отношения R_1/R_2. Действительно, хотя величина R_2 и входит в правую часть (25.54), она всегда является коэффициентом при разности $\theta - \theta_0$, которая изменяется от 0 до ∞, так что $R_2(\theta - \theta_0)$ также пробегает значения от 0 до ∞, каково бы ни было R_2. Обозначая в (25.54) $\Delta = R_2(\theta - \theta_0)$, получаем, следовательно,

$$d(\Delta) = G\{\Delta - \lambda_\alpha\} - G\left\{\Delta \frac{R_1}{R_2} - \lambda_\alpha\right\}. \qquad (25.55)$$

Первая производная от (25.55) по Δ равна

$$d'(\Delta) = g\{\Delta - \lambda_\alpha\} - \frac{R_1}{R_2} g\left\{\Delta \frac{R_1}{R_2} - \lambda_\alpha\right\}.$$

Приравнивая ее к нулю, находим

$$\frac{R_1}{R_2} = \frac{g\{\Delta - \lambda_\alpha\}}{g\left\{\Delta \frac{R_1}{R_2} - \lambda_\alpha\right\}} = \exp\left\{-\frac{1}{2}(\Delta - \lambda_\alpha)^2 + \frac{1}{2}\left(\Delta \frac{R_1}{R_2} - \lambda_\alpha\right)^2\right\} = \\ = \exp\left\{-\frac{1}{2}\Delta^2\left(1 - \frac{R_1^2}{R_2^2}\right) + \lambda_\alpha \Delta\left(1 - \frac{R_1}{R_2}\right)\right\}. \qquad (25.56)$$

Соотношение (25.56) является квадратным уравнением относительно Δ, положительный корень которого равен

$$\Delta = \frac{\lambda_\alpha + \left\{\lambda_\alpha^2 + 2\left(\frac{1 + (R_1/R_2)}{1 - (R_1/R_2)}\right)\log\frac{R_2}{R_1}\right\}^{1/2}}{1 + (R_1/R_2)}. \qquad (25.57)$$

При этом значении функция (25.55) достигает максимума. Рассмотрим, например, случай $\alpha = 0{,}05$ ($\lambda_\alpha = 1{,}645$) и $R_1/R_2 = 0{,}5$. Тогда (25.57) дает

$$\Delta = \frac{1{,}645 + \{1{,}645^2 + 6 \log_e 2\}^{1/2}}{1{,}5} = 2{,}85.$$

С помощью таблиц нормальной ф. р. имеем тогда в (25.55)

$$P_2 = G\{2{,}85 - 1{,}64\} = G\{1{,}21\} = 0{,}89,$$
$$P_1 = G\{1{,}42 - 1{,}64\} = G\{-0{,}22\} = 0{,}41.$$

С помощью графического метода, эквивалентного приведенному выше, Д. Кокс и Стьюарт (1955) нашли значения P_2 и P_1 в точке максимальной разности для ряда значений α и R_1/R_2. Их таблица приведена ниже; в ней содержится и рассмотренный нами пример.

Асимптотические мощности критериев в точке максимальной разности (Д. Кокс и Стьюарт, 1955) (значения P_i даны в процентах)

R_1/R_2 \ α	0,10		0,05		0,01		0,001	
	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
0,9	67	73	63	71	49	60	54	67
0,8	61	74	56	72	49	71	43	72
0,7	59	80	51	77	42	77	39	83
0,6	54	84	47	84	39	86	29	87
0,5	48	88	41	89	30	90	20	93
0,3	35	96	27	96	14	97	07	99

Из таблицы видно, что, когда α убывает при фиксированной величине R_1/R_2, максимальная разность между асимптотическими функциями мощности возрастает. В действительности путем выбора достаточно малого α ее можно сделать как угодно близкой к 1. Аналогично, при фиксированном α максимальная разность возрастает с убыванием R_1/R_2.

Из таблицы мы получаем полезное для практики следствие. Если R_1/R_2 равно или больше, чем 0,9, то потеря мощности *вдоль всей* асимптотической функции мощности не превосходит 0,08 при $\alpha = 0{,}05$; 0,11 при $\alpha = 0{,}01$ и 0,13 при $\alpha = 0{,}001$, наиболее часто используемых размерах критерия. Поскольку согласно (25.36) R_1/R_2 равно отношению первых производных функций мощности, то из (25.37) получаем, что $(R_1/R_2)^{1/\delta} = A_{12}$, где δ обычно равно 1/2. Таким образом, АОЭ должна равняться $(0{,}9)^\delta$ для того, чтобы приведенные выше утверждения были справедливы.

АОЭ и эффективность оценивания

25.13 Существует простая связь между АОЭ и эффективностью оценивания. Пусть, как и прежде, имеется две состоятельных статистики критериев t_i. Определим не зависящие от n функции f_i так, чтобы статистики

$$T_i = f_i(t_i) \tag{25.58}$$

были состоятельными *оценками* для θ. Если обозначим

$$\theta = f_i(\tau_i), \tag{25.59}$$

то из (25.58) будет следовать, что (поскольку $T_i \to \theta$ по вероятности) $t_i \to \tau_i$ и $\mathbf{M}(t_i)$ (если оно существует) также стремится к τ_i. Разлагая (25.58) по теореме Тейлора в окрестности точки τ_i и используя (25.59), имеем

$$T_i = \theta + (t_i - \tau_i)\left[\frac{\partial f(t_i)}{\partial \tau_i}\right]_{t_i = t_i^*}, \tag{25.60}$$

где t_i^* — промежуточное значение между t_i и τ_i, стремящееся к τ_i с ростом n. Таким образом, (25.60) может быть переписано как

$$T_i - \theta \sim (t_i - \tau_i)\left[\frac{\partial \theta}{\partial \mathbf{M}(t_i)}\right],$$

откуда

$$\mathbf{D}T_i \sim \mathbf{D}t_i \Big/ \left(\frac{\partial \mathbf{M}(t_i)}{\partial \theta}\right)^2. \tag{25.61}$$

Если 2δ есть порядок малости по n дисперсий статистик T_i, то согласно (17.65) и (25.61) эффективность оценки T_2 по сравнению с T_1 равна

$$\lim_{n\to\infty}\left(\frac{\mathbf{D}T_1}{\mathbf{D}T_2}\right)^{1/(2\delta)} = \left[\frac{\{\partial \mathbf{M}(t_2)/\partial\theta\}^2/\mathbf{D}t_2}{\{\partial \mathbf{M}(t_1)/\partial\theta\}^2/\mathbf{D}t_1}\right]^{1/(2\delta)}. \tag{25.62}$$

Если $m_i = 1$, то в точке θ_0 (25.62) равно в точности АОЭ (25.27). Итак, АОЭ, по существу, совпадает с относительными эффективностями тех преобразований статистик критериев, которые являются состоятельными оценками рассматриваемых параметров. Но это соответствие является локальным: в **22.15** мы видели, что в общем случае нет строгой связи между эффективностью оценки и мощностью. Легко видеть, что критерии, основанные на эффективных оценках, имеют максимальную АОЭ и (согласно **25.8—11**) максимальную производную функции мощности в точке θ_0. Из (25.62) и (17.61) вытекает также, что если статистика T_1 эффективна, то $A_{21} = \{\rho(T_1, T_2)\}^{1/\delta}$. Более общий результат в терминах $\rho(t_1, t_2)$ приведен в упражнении 25.9.

Пример 25.5

Полученный только что результат объясняет тот факт, отмеченный в примере 25.2, что АОЭ выборочной медианы по сравнению с выборочным средним при проверке гипотезы о среднем значении нормальной совокупности равна эффективности медианы как оценки для этого параметра.

Случаи не нормальных распределений

25.14 Начиная с **25.5**, мы ограничились случаем, когда статистики критериев асимптотически нормальны. Однако исследование пунктов **25.5—7** показывает, что при выводе АОЭ не использовалась специфика нормального распределения. Мы устанавливали условия, при которых аргументы u_i функции мощности $G\{u_i\}$ в (25.19) были бы равны между собой на последовательности альтернатив (25.17). Роль функции G сводилась к тому, чтобы обеспечить одинаковую форму асимптотических функций мощности, и нам требовалось лишь, чтобы она была достаточно регулярной ф. р.

Следовательно, если асимптотические функции мощности двух критериев даются любой двухпараметрической функцией G, причем только один из параметров зависит от θ, то результаты пунктов **25.5—7** останутся в силе, так как (25.17) будет фиксировать этот параметр, а величина u_i в (25.19) определяет тогда другой параметр. При заданной форме G критическая область одностороннего критерия всегда может быть представлена в виде (25.12), где λ_α интерпретируется как множитель перед D_{i0}, требуемый для того, чтобы сделать (25.12) критической областью размера α.

25.15 Единственным важным предельным распределением, отличным от нормального, является нецентральное χ^2-распределение, свойства которого обсуждались в **24.4—5**. Предположим, что для проверки гипотезы H_0: $\theta = \theta_0$ имеются две статистики критериев t_i с такими распределениями, имеющими числа степеней свободы ν_i (не зависящие от θ) и параметры нецентральности $\lambda_i(\theta)$, где $\lambda_i(\theta_0) = 0$, так что, когда выполняется H_0, χ^2-распределения становятся центральными. Имеем (см. упражнение 24.1)

$$E_{i1} = \nu_i + \lambda_i(\theta), \qquad D_{i0}^2 = 2\nu_i. \tag{25.63}$$

Следовательно, все результаты из **25.5—6**, относящиеся к односторонним критериям, справедливы для сравнения статистик критериев, распределенных как нецентральные χ^2-величины (центральные, если выполняется H_0) с числом степеней

свободы, не зависящим от θ. В частности, когда $\delta_1 = \delta_2 = \delta$ и $m_1 = m_2 = m$, подстановка (25.63) в (25.27) дает

$$A_{21} = \lim_{n\to\infty} \left\{ \frac{\lambda_2^{(m)}(\theta_0)/\nu_2^{1/2}}{\lambda_1^{(m)}(\theta_0)/\nu_1^{1/2}} \right\}^{1/(m\delta)}. \tag{25.64}$$

Иной вывод этого результата приводит Э. Хеннан (1956).

Другие меры эффективности критериев

25.16 Хотя в последующих главах в качестве мер эффективности критериев будут использоваться только относительная эффективность и АОЭ, мы закончим эту главу обсуждением двух других предлагавшихся методов.

Уолш (1946) предложил сравнивать критерии при фиксированном размере α на основе меры, которая принимает в расчет качество критериев при всех альтернативных значениях параметра θ. Если критерии t_i основаны на выборках объемов n_i и имеют функции мощности $P_i(\theta, n_i)$, то эффективность t_2 по сравнению с t_1 есть $n_1/n_2 = e_{12}$, где

$$\int [P_1(\theta, n_1) - P_2(\theta, n_2)]\, d\theta = 0. \tag{25.65}$$

Таким образом, при заданном объеме одной из выборок (скажем, n_2) мы выбираем n_1 так, чтобы алгебраическая сумма площадей между функциями мощности равнялась нулю. Тогда мера эффективности есть n_1/n_2.

Эта мера разумным образом устраняет влияние θ из таблицы с тремя входами, требуемой для сравнения двух функций мощности. Однако e_{12} по-прежнему зависит от α и, что более важно, от n_2. Кроме того, вычисление такого n_1/n_2, чтобы удовлетворялось соотношение (25.65), трудоемко, и, возможно, вследствие этого приведенная мера редко используется. Тем не менее в качестве асимптотической меры она эквивалентна АОЭ, по крайней мере для асимптотически нормальных статистик критериев с $m_i = 1$ в (25.15). Действительно, как и в **25.12**, имеем

$$P_i(\theta, n_i) = G\{(\theta - \theta_0) R_i - \lambda_\alpha\},$$

где $R_i = E_i'(\theta_0)/D_{i0}$ (как в (25.16)), и тогда (25.65) принимает вид

$$\int [G\{(\theta - \theta_0) R_1 - \lambda_\alpha\} - G\{(\theta - \theta_0) R_2 - \lambda_\alpha\}]\, d\theta = 0. \tag{25.66}$$

Ясно, что (25.66) выполняется асимптотически только тогда,

когда $R_1 = R_2$ или, согласно (25.16), когда

$$\frac{R_1}{R_2} \sim \frac{c_1}{c_2}\left(\frac{n_1}{n_2}\right)^{\delta} = 1,$$

откуда

$$\lim \frac{n_1}{n_2} = \left(\frac{c_2}{c_1}\right)^{1/\delta} = A_{21},$$

что соответствует (25.27) с $m = 1$.

25.17 В заключение приведем принадлежащий Чернову (1952) совершенно иной подход к задаче измерения асимптотической эффективности критериев. Для случайной величины x с производящей функцией моментов $\mathbf{M}_x(t) = \mathbf{M}(e^{xt})$ определим

$$m(a) = \inf_t \mathbf{M}_{x-a}(t) \tag{25.67}$$

— абсолютно минимальное значение п. ф. м. величины $(x - a)$. Если $\mathbf{M}(x \mid H_i) = \mu_i$ для простых гипотез H_0, H_1, то положим

$$\rho = \inf_{\mu_0 \leqslant a \leqslant \mu_1} \max\{m_0(a),\ m_1(a)\}, \tag{25.68}$$

где индекс у функций m, определенных в (25.67), указывает на соответствующую гипотезу. Пусть односторонний критерий для гипотезы H_0 против H_1 основан на сумме n одинаково распределенных величин x_i и имеет уровень α и мощность $1 - \beta$. Для такого критерия Чернов показал, что если минимизировать любую линейную функцию $l(\alpha, \beta)$ от вероятностей ошибок α и β, то ее минимальное значение при $n \to \infty$ ведет себя, как ρ^n, где ρ определено в (25.68). Рассмотрим два таких критерия t_i, основанных на выборках объемов n_i. Если они имеют одинаковые минимумы $l(\alpha, \beta)$, то

$$\lim \frac{\rho_1^{n_1}}{\rho_2^{n_2}} = 1,$$

или

$$\lim \frac{n_1}{n_2} = \frac{\log \rho_2}{\log \rho_1}. \tag{25.69}$$

Таким образом, правая часть (25.69) является мерой асимптотической эффективности критерия t_2 по сравнению с t_1. Ее использование ограничено случаями, в которых статистики критериев основаны на суммах независимых наблюдений, и при этом могут потребоваться значительные вычисления.

25.18 Гёфдинг (1965) развил метод сравнения критериев, когда $\alpha \to 0$ при $n \to \infty$ (в отличие от подхода, используемого в 25.5 и далее, где α фиксировано, а $H_1 \to H_0$ при $n \to \infty$), и, пользуясь этим методом, показал, что критерии ОП обладают оптимальным свойством для гипотез о мультиномиальных распределениях.

УПРАЖНЕНИЯ

25.1 Пусть проверяется гипотеза H_0 о том, что медиана генеральной совокупности имеет заданное значение θ_0. Критерий знаков состоит в том, что подсчитывается число наблюдений, превысивших θ_0, и H_0 отвергается, когда это число слишком велико. Показать, что для нормальной совокупности этот критерий имеет АОЭ, равную $2/\pi$, по сравнению с t-критерием Стьюдента (для H_0), и связать это с результатом примера 25.2.

(Кокрэн, 1937.)

25.2 Обобщая результат упражнения 25.1, показать, что для любой непрерывной функции плотности f с дисперсией σ^2 АОЭ критерия знаков по сравнению с t-критерием равна $4\sigma^2\{f(\theta_0)\}^2$.

(Питмэн, 1948.)

25.3 Гипотеза о разности между средними двух нормальных совокупностей с одинаковыми дисперсиями проверяется на основе двух независимых выборок путем сравнения каждого наблюдения y_j из второй выборки с каждым наблюдением x_i из первой выборки и подсчета, сколько раз y_j превзойдет x_i. Показать, что АОЭ этого S-критерия по сравнению с двухвыборочным t-критерием Стьюдента равна $3/\pi$.

(Питмэн, 1948.)

25.4 Обобщая результат упражнения 25.3, показать, что если любые две непрерывные функции плотности $f(x)$, $f(x-a)$ отличаются только параметром расположения и имеют дисперсию σ^2, то АОЭ S-критерия по сравнению с t-критерием равна

$$12\sigma^2\left\{\int_{-\infty}^{\infty}\{f(x)\}^2\,dx\right\}^2.$$

(Питмэн, 1948.)

25.5 Пусть x — нормальная величина со средним μ_i и дисперсией σ_i^2 при гипотезе H_i $(i=0,\ 1;\ \mu_0<\mu_1)$. Показать, что величина (25.68) равна

$$\rho=\exp\left\{-\frac{1}{2}\left(\frac{\mu_1-\mu_0}{\sigma_1+\sigma_0}\right)^2\right\}.$$

(Чернов, 1952.)

25.6 Пусть x/σ_i^2 имеет хи-квадрат распределение с r степенями свободы при гипотезе H_i $\left(i=0,\ 1;\ \sigma_0^2/\sigma_1^2=\tau<1\right)$. Показать, что ρ из (25.68) удовлетворяет соотношению

$$\log\rho=-\frac{1}{2}r\,(\delta-1-\log\delta),$$

где

$$\delta=(\log\tau)/(\tau-1).$$

(Чернов, 1952.)

25.7 Пусть t_1 и t_2 — несмещенные оценки параметра θ, имеющие при больших выборках совместно нормальное распределение с дисперсиями σ^2, σ^2/e соответственно $(0<e\leqslant 1)$. Используя результаты из **16.23** и **17.29**, показать, что

$$\mathbf{M}\,(t_1\mid t_2)=\theta\,(1-e)+t_2e$$

и, следовательно, если наблюденное значение t_2 отличается от θ на d своих стандартных отклонений, то ожидаемое значение оценки t_1 будет отличаться от θ на $de^{1/2}$ ее стандартных отклонений.

(Д. Кокс, 1956.)

25.8 Используя упражнение 25.7, показать, что если t_2 применяется для проверки гипотезы H_0: $\theta = \theta_0$, то мы можем вычислить «ожидаемый результат» для критерия, основанного на более эффективной статистике t_1. В частности, показать, что если односторонний критерий размера 0,01, использующий t_2, отвергает H_0, то мы можем ожидать, что односторонний критерий размера 0,05, использующий t_1, также будет отвергать H_0, если $e > 0{,}50$; если же критерий с «равными хвостами» размера 0,01, основанный на t_2, отвергает H_0, то можно ожидать, что критерий с «равными хвостами» уровня 0,05, основанный на t_1, будет отвергать H_0, если $e > 0{,}58$.

25.9 Пусть t_1 — статистика с максимальной АОЭ и t_2 — некоторая другая статистика критерия для той же задачи с δ и m такими же, как у t_1. Рассматривая взвешенное среднее

$$t_3 = a\frac{t_1}{D_{10}} + (1-a)\frac{t_2}{D_{20}},$$

показать, что в (25.16)

$$R_3 = \{aR_1 + (1-a)R_2\}/\{a^2 + (1-a)^2 + 2a(1-a)\rho\}^{1/2},$$

где ρ — асимптотический коэффициент корреляции между t_1 и t_2. Вывести отсюда, что ρ равен R_2/R_1, т. е. равен АОЭ (25.27) в степени $m\delta$.

(См. Ван Эден, 1963.)

ГЛАВА 26

СТАТИСТИЧЕСКАЯ ЗАВИСИМОСТЬ: ЛИНЕЙНАЯ РЕГРЕССИЯ И КОРРЕЛЯЦИЯ

26.1 В этой и трех последующих главах мы будем интересоваться теми или иными аспектами зависимостей между двумя или большим числом величин. Ранее уже обсуждались двумерные и многомерные распределения, их моменты и семиинварианты: в частности, исследовались свойства двумерных и многомерных нормальных распределений. Однако систематическое изучение зависимостей между величинами было отложено до завершения изложения теории оценивания и проверки гипотез. Но даже этих четырех глав нам недостаточно, чтобы целиком охватить всю проблему. Поэтому более сложные вопросы, связанные с распределением трех и более величин, мы отложим до третьего тома, где будет рассмотрен многомерный анализ.

26.2 Так как нам предстоит исследовать очень обширную тему, то полезно начать с общего обзора.

Бóльшая часть работы по этой теме возникла в связи с задачей о совместном распределении пары случайных величин; ее можно назвать задачей о *статистической зависимости*. Существует совершенно отличная область, касающаяся зависимостей строго функционального вида между величинами (как, например, зависимости в классической физике). Указанный вид зависимостей тоже представляет статистический интерес, потому что функционально связанные величины подвержены ошибкам наблюдений или измерений. Назовем это задачей о *функциональной зависимости* и отложим ее изучение до главы 29. А до тех пор мы будем заниматься только задачей о статистической зависимости, в которой величины (кроме вырожденных случаев) не связаны функционально, хотя могут быть подвержены ошибкам наблюдений и измерений. Мы будем рассматривать их просто как совокупность случайных величин, подчиненных некоторому совместному распределению.

26.3 В самой области статистической зависимости полезно провести дальнейшее различие. Нас может интересовать либо *взаимозависимость* между несколькими величинами (не обязательно между всеми), либо *зависимость* одной или большего числа величин от остальных. Например, можно рассмотреть во-

прос, существует ли у людей связь между длиной руки и длиной ноги; при такой постановке это есть задача о взаимозависимости. Но если мы хотим, используя измерения длин ног, получить информацию о длине рук, то мы приходим к задаче о зависимости длины руки от длины ноги. Это пример ситуации, в которой может представлять интерес как взаимозависимость, так и зависимость. С другой стороны, имеются ситуации, в которых интересна только зависимость. Связь между величиной урожая и количеством выпавших осадков представляет собой пример существенной асимметрии. Здесь из внестатических соображений понятно, что дожди влияют на урожай и, совершенно определенно, урожай не воздействует на дожди. Таким образом, мы должны изучать зависимость урожая от дождей.

Не существует четкого различия в статистической терминологии для этих существенно разных типов задач. Например, в главе 27 мы увидим, что если нас интересует взаимозависимость двух величин, когда устранено воздействие остальных величин, то мы используем метод так называемой «частной корреляции». Рассматривая же зависимость единственной величины от группы других, мы применяем «множественную корреляцию». Тем не менее в основном верно, что исследование *взаимозависимости* приводит к теории корреляции, содержащейся в главах 26, 27, тогда как изучение *зависимости* ведет к теории регрессии, обсуждаемой в тех же главах, а также в главе 28.

26.4 Прежде чем перейти к изложению теории корреляции (развитой в основном в конце прошлого и в начале этого столетий Карлом Пирсоном и Юлом), которая займет бóльшую часть данной главы, сделаем одно заключительное общее замечание. Статистическая зависимость, как бы ни была она сильна, никогда не может установить *причинной* связи: наши идеи о причине должны приходить извне статистики, в конечном счете из некоторой другой теории. Даже в простом примере о размере урожая и количестве осадков из **26.3** мы не имеем *статистических* причин для отказа от идеи зависимости дождей от урожая: отказ сделан на основе совершенно других соображений. И даже если бы дожди и урожай были в полном функциональном соответствии, то мы все равно не подумали бы обратить эту «очевидную» причинную связь. Нам нет нужды углубляться в философское обсуждение этого вопроса; для наших целей необходимо только еще раз подчеркнуть, что статистическая зависимость любого сорта логически не влечет причинной.

Бернард Шоу блестяще сказал об этом в своем предисловии к «Доктору на распутье» (1906): «Даже опытные статистики часто оказываются не в состоянии оценить, до какой степени смысл статистических данных искажается молчаливыми пред-

положениями их интерпретаторов... Легко доказать, что ношение цилиндров и зонтиков расширяет грудную клетку, удлиняет жизнь и дает относительный иммунитет от болезней... Университетский диплом, ежедневная ванна, обладание тридцатью парами брюк, знание музыки Вагнера, скамья в церкви — короче, все, что подразумевает большие средства и хорошее воспитание, ...может быть с помощью статистики представлено как магические чары, дарующие привилегии любого сорта. Математик, чьи корреляции привели бы в восхищение Ньютона, может, собирая данные и делая из них выводы, впасть в совершенно грубые ошибки на основе таких же популярных заблуждений, как описанные выше».

Хотя Шоу в данном случае отстаивает характерно сомнительную точку зрения, его логика обоснована. Последователи Карла Пирсона и Юла в первом приступе энтузиазма, порожденного корреляционной техникой, легко делали опрометчивые выводы. Это продолжалось до тех пор, пока (спустя двадцать лет после написанного Шоу) Юл (1926) не напугал статистиков примерами высоких корреляций, которые, очевидно, не выражали причинных связей: например, количество самоубийств за год было сильно коррелировано с принадлежностью к англиканской церкви. Большинство этих «бессмысленных» корреляций действует через сопутствующие изменения во времени. Упомянутые примеры имели благотворный эффект, доводя до сознания статистиков, что причинная зависимость не может быть выведена ни из какого наблюдаемого совместного изменения, даже самого тесного. Теперь, спустя более чем тридцать лет, статистики впали в другую крайность: корреляционный анализ стал совершенно немодным. Имеется, однако, широкое поле приложений (например, социальные науки и психология), где характеры причин еще недостаточно хорошо поняты для того, чтобы корреляционный анализ был заменен более специфическими «структурными» статистическими методами. Есть, кроме того, обширная область многомерного анализа, где вычисление и исследование матрицы коэффициентов корреляции является необходимой прелюдией к детальному статистическому анализу. Все это делает необходимым изучение нашего предмета.

26.5 В главе 1 (таблицы 1.15, 1.23, 1.24) было дано несколько примеров двумерных распределений, возникающих в практике. Таблицы 26.1 и 26.2 содержат еще два примера, которые будут использованы в качестве иллюстраций.

Пока будем рассматривать эти данные как генеральные совокупности, оставляя на дальнейшее вопрос об их выборе. Так же как для одномерных распределений мы строили результирующие константы, такие, как среднее, дисперсия и др., нам

Таблица 26.1

Распределение веса и роста 4995 женщин Великобритании, 1951. (Воспроизведено, по разрешению, из *Women's Measurements and Sizes*, London, H. M. S. O., 1957)

Вес (y): центральные значения интервалов в фунтах	Рост (x): центральные значения интервалов в дюймах											Всего
	54	56	58	60	62	64	66	68	70	72	74	
278,5						1						1
272,5												—
266,5						1						1
260,5							1					1
254,5												—
248,5					1	1						2
242,5							1					1
236,5							1					1
230,5					2				1			3
224,5					1	2	1					4
218,5			1		2	1		1				5
212,5				2	1	6		1	1			11
206,5				2	2	3	2		1			10
200,5			4	2	6	2						14
194,5				1	3	7	7	4	1			23
188,5			1	5	14	8	12	3	1	2		46
182,5			1	7	12	26	9	5		1	2	63
176,5			5	8	18	21	15	11	7		2	87
170,5			2	11	17	44	21	13	3	1		112
164,5		1	3	12	35	48	30	15	5	3		152
158,5			8	17	52	42	36	21	9			185
152,5		1	7	30	81	71	58	21	2	2		273
146,5		2	13	36	76	91	82	36	8	1		345
140,5		1	6	55	101	138	89	50	8			448
134,5			15	64	95	175	122	45	5			521
128,5		1	19	73	155	207	101	25	3			584
122,5		3	34	91	168	200	81	12	1	1		591
116,5		3	24	108	184	184	50	8				561
110,5		5	33	119	165	124	22	4				472
104,5	1	3	33	87	95	35	6					260
98,5	2	5	29	59	45	16	3					159
92,5		6	10	21	9							46
86,5		1	5	3								9
80,5	2	1	1									4
Всего	5	33	254	813	1340	1454	750	275	56	11	4	4995

хотелось бы теперь выразить связь между величинами и, в частности, их взаимозависимость. Результирующие константы для двумерного распределения естественно возникают из следующих рассуждений.

Таблица 26.2

Распределение объема груди и роста для 4995 женщин Великобритании, 1951. (Данные из того же источника, что и таблица 26.1)

Объем груди (y): центральные значения интервалов в дюймах	Рост (x): центральные значения интервалов в дюймах											Всего
	54	56	58	60	62	64	66	68	70	72	74	
56					1							1
54					1	2						3
52			1		3	4	1		1			10
50			1	3	5	4	1					14
48		1	3	9	7	6	3	1				30
46			4	11	17	17	7		1			57
44		2	11	26	50	45	17	10	1			162
42		2	11	42	85	73	31	12	3	2		261
40		2	20	76	132	131	71	31	9	4	3	479
38		2	36	98	158	203	126	65	17	3	1	709
36		6	48	188	317	410	263	89	15	1		1337
34	1	9	67	210	376	427	196	59	8			1353
32	3	5	39	131	163	122	31	8	1	1		504
30	1	4	11	18	25	10	2					71
28			2	1			1					4
Всего	5	33	254	813	1340	1454	750	275	56	11	4	4995

Обозначим две случайные величины через x, y. Для любого заданного значения величины x, например X, распределение величины y называется y-сечением. Разумеется, y-сечение представляет собой условное распределение y при заданном $x = X$. Это условное распределение имеет среднее, которое обозначим

$$\bar{y}_X = \mathbf{M}(y \mid X) \tag{26.1}$$

и которое является функцией от X. Аналогично, рассматривая x-сечение для $y = Y$, находим

$$\bar{x}_Y = \mathbf{M}(x \mid Y). \tag{26.2}$$

Равенства (26.1) и (26.2) определяют *линии регрессии* (или, короче, *регрессии*) y по x и x по y соответственно.

В дальнейшем мы не будем использовать заглавную букву для обозначения величины, значение которой фиксировано, если из контекста понятны обозначения $\mathbf{M}(y \mid x)$, $\mathbf{M}(x \mid y)$.

На рис. 26.1 и 26.2, построенных по данным таблиц 26.1 и 26.2, нанесены средние для y-сечений (крестики) и для x-сечений (кружочки). Линии CC' и RR' проведены так, чтобы как можно лучше (в смысле наименьших квадратов, см. **26.8**)

приблизить средние сечений с помощью прямых линий. Эти диаграммы характеризуют двумерное распределение подобно тому, как среднее характеризует одномерное. Поскольку данные группированы, то у нас есть информация только относительно сечений, соответствующих группированным значениям x-ов и y-ов. Мы уславливаемся сопоставлять y-сечения центральным

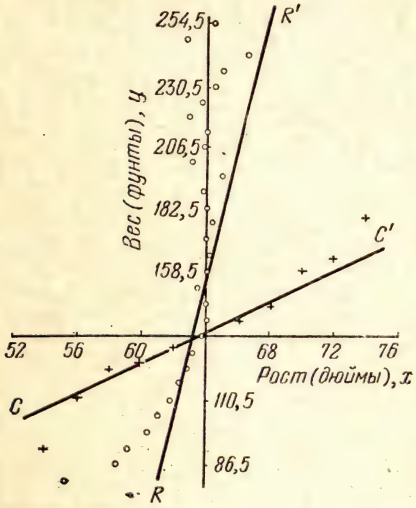


Рис. 26.1. Линии регрессии для данных таблицы 26.1.

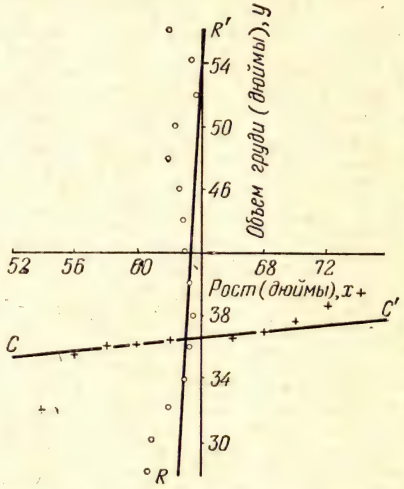


Рис. 26.2. Линии регрессии для данных таблицы 26.2.

значениям x-интервалов, в пределах которых они наблюдались, и аналогично для x-сечений.

Непосредственно регрессия будет изучаться в главе 28. Здесь же мы используем понятие линейной регрессии главным образом для того, чтобы ввести некоторую меру взаимозависимости, а именно коэффициент корреляции (смешанный момент). Кроме того, мы воспользуемся случаем дополнить наше изучение двумерного нормального распределения с точки зрения регрессии и корреляции.

Ковариация и регрессия

26.6 Кажется естественным использовать в качестве основы для меры зависимости смешанный момент μ_{11}, с которым мы уже несколько раз встречались в предыдущих главах. Момент μ_{11}, называемый *ковариацией* между x и y, в случае дискретной совокупности определяется соотношением

$$\mu_{11} = \sum_{i=1}^{n} (x_i - \mu_x)(y_i - \mu_y)/n \equiv \sum_{i=1}^{n} x_i y_i/n - \mu_x \mu_y, \qquad (26.3)$$

где n — число пар значений x, y, а μ_x, μ_y — средние величин x, y. В случае непрерывной совокупности с распределением

$$dF(x, y) = f(x, y)\,dx\,dy \tag{26.4}$$

соответствующая формула имеет вид (см. **3.27**)

$$\mu_{11} = \int\int (x - \mu_x)(y - \mu_y)\,dF(x, y) = \\ = \mathbf{M}\{(x - \mu_x)(y - \mu_y)\} \equiv \mathbf{M}(xy) - \mathbf{M}(x)\,\mathbf{M}(y). \tag{26.5}$$

Если величины x, y независимы, то, как мы видели в примере 12.7,

$$\mu_{11} = \varkappa_{11} = 0. \tag{26.6}$$

Согласно этому же примеру обратное, вообще говоря, неверно: из (26.6) в общем случае не следует независимость, для которой требуется, чтобы

$$\varkappa_{rs} = 0 \quad \text{при всех} \quad r \times s \neq 0. \tag{26.7}$$

Однако в случае двумерного *нормального* распределения $\varkappa_{rs} = 0$ при всех $r + s > 2$, так что $\varkappa_{11}$ есть единственный ненулевой смешанный семиинвариант. Таким образом, (26.6) влечет (26.7) и независимость для нормальных величин. Это может оказаться справедливым и для других конкретных распределений, но в общем случае это не так. Пример 26.1 содержит распределение, отличное от нормального, для которого равенство $\varkappa_{11} = 0$ влечет независимость; в примере же 26.2 приведено распределение, для которого это неверно.

Пример 26.1

Пусть x и y — нормированные величины, имеющие двумерное нормальное распределение. Тогда совместная характеристическая функция величин x^2 и y^2 равна

$$\varphi(t, u) = \frac{1}{2\pi(1 - \rho^2)^{1/2}} \int\int \exp\left\{ -\frac{1}{2(1 - \rho^2)} [x^2\{1 - 2(1 - \rho^2)\,it\} - \right. \\ \left. - 2\rho xy + y^2\{1 - 2(1 - \rho^2)\,iu\}] \right\} dx\,dy.$$

Согласно упражнению 1.5 двойной интеграл равен

$$\pi \cdot 2(1 - \rho^2) \begin{vmatrix} 1 - 2(1 - \rho^2)\,it & \rho \\ \rho & 1 - 2(1 - \rho^2)\,iu \end{vmatrix}^{-1/2}.$$

Поэтому

$$\varphi(t, u) = (1 - \rho^2)^{1/2} [\{1 - 2(1 - \rho^2)\,it\}\{1 - 2(1 - \rho^2)\,iu\} - \rho^2]^{-1/2} = \\ = [(1 - 2it)(1 - 2iu) + 4\rho^2 tu]^{-1/2}.$$

Отсюда получаем

$$\varphi(t, 0) = (1 - 2it)^{-1/2}, \quad \varphi(0, u) = (1 - 2iu)^{-1/2},$$

так что маргинальные распределения являются хи-квадрат распределениями с одной степенью свободы, что нам уже известно. Дифференцируя логарифм функции $\varphi(t, u)$, находим

$$\mu_{11} = \varkappa_{11} = \left[\frac{\partial^2 \log \varphi(t, u)}{\partial (it)\, \partial (iu)}\right]_{t=u=0} = 2\rho^2.$$

Когда $\rho = 0$, мы имеем

$$\varphi(t, u) = \varphi(t, 0)\, \varphi(0, u).$$

Согласно **4.16—17** это равенство служит необходимым и достаточным условием для независимости x^2 и y^2. Следовательно, $\mu_{11} = 0$ влечет в данном случае независимость.

Пример 26.2

Пусть величины x и y имеют двумерное распределение, равномерное на круге радиуса единица с центром в начале координат. Имеем

$$dF(x, y) = dx\, dy/\pi, \qquad 0 \leqslant x^2 + y^2 \leqslant 1,$$

откуда

$$\mu_{11} = \int\int xy\, dF = \frac{1}{\pi} \int\int xy\, dx\, dy =$$

$$= \frac{1}{\pi} \int x \left[\int y\, dy\right] dx = \frac{1}{\pi} \int x \left[\frac{1}{2} y^2\right]_{-(1-x^2)^{1/2}}^{+(1-x^2)^{1/2}} dx = 0,$$

что, впрочем, очевидно. Но поскольку интервал изменения каждой из величин зависит от значения другой, то величины x и y не являются независимыми.

Линейная регрессия

26.7 Если регрессия x по y, определенная в (26.2), в точности линейна, то имеет место уравнение

$$\mathbf{M}(x \mid y) = \alpha_1 + \beta_1 y, \tag{26.8}$$

для которого сейчас найдем коэффициенты α_1 и β_1. Взяв от обеих частей (26.8) математическое ожидание, получим

$$\mu_x = \alpha_1 + \beta_1 \mu_y. \tag{26.9}$$

Вычитая (26.9) из (26.8), умножая обе части на $(y-\mu_y)$ и беря вновь математические ожидания, получаем

$$\mathsf{M}\{(x-\mu_x)(y-\mu_y)\}=\beta_1\mathsf{M}\{(y-\mu_y)^2\},$$

или

$$\beta_1=\mu_{11}/\sigma_2^2, \tag{26.10}$$

где σ_2^2 — дисперсия y. Аналогично, из (26.1) находим

$$\beta_2=\mu_{11}/\sigma_1^2 \tag{26.11}$$

для коэффициента в точной линейной регрессии y по x. Соотношения (26.10) и (26.11) определяют (линейные) *коэффициенты регрессии**) x по y (β_1) и y по x (β_2). Используя (26.8), (26.9) и (26.10), находим

$$\mathsf{M}(x\mid y)-\mu_x=\beta_1(y-\mu_y) \tag{26.12}$$

и, аналогично,

$$\mathsf{M}(y\mid x)-\mu_y=\beta_2(x-\mu_x). \tag{26.13}$$

Равенства (26.12) и (26.13) называются *уравнениями линейной регрессии.*

Со случаем точной линейной регрессии мы уже встречались при рассмотрении двумерного нормального распределения в **16.23**.

Пример 26.3

В примере 26.1 регрессии x^2 по y^2 и y^2 по x^2 строго линейны. Действительно, из **16.23** при $\sigma_1=\sigma_2=1$ в (16.46) имеем

$$\mathsf{M}(x\mid y)=\rho y,\qquad \mathsf{D}(x\mid y)=1-\rho^2.$$

Таким образом,

$$\mathsf{M}(x^2\mid y)\equiv\mathsf{D}(x\mid y)+\{\mathsf{M}(x\mid y)\}^2=1-\rho^2+\rho^2y^2.$$

Каждому y^2 соответствуют значения $+y$ и $-y$, появляющиеся с равными вероятностями. И поскольку $\mathsf{M}(x^2\mid y)$ есть функция только от y^2, то

$$\mathsf{M}(x^2\mid y^2)=\frac{1}{2}\{\mathsf{M}(x^2\mid y)+\mathsf{M}(x^2\mid -y)\}=\mathsf{M}(x^2\mid y)=1-\rho^2+\rho^2y^2,$$

что можно переписать в виде (26.12):

$$\mathsf{M}(x^2\mid y^2)-1=\rho^2(y^2-1).$$

Следовательно, регрессия x^2 по y^2 строго линейна. Аналогично,

$$\mathsf{M}(y^2\mid x^2)-1=\rho^2(x^2-1).$$

---

*) Обозначения β_1, β_2 не связаны с обозначениями асимметрии и эксцесса в **3.31—32**; они вряд ли могут быть перепутаны, так как встречаются в разных контекстах.

В примере 26.1 мы видели, что $\mu_{11} = 2\rho^2$, а дисперсия хи-квадрат распределения с одной степенью свободы, как известно, равна 2. Поэтому равенства (26.10) и (26.11) подтверждают, что ρ^2 является коэффициентом регрессии в каждом из уравнений линейной регрессии.

Пример 26.4

В условиях примера 26.2 легко видеть, что $\mathbf{M}(x|y) = \mathbf{M}(y|x) = 0$. Следовательно, в данном случае получаем линейные регрессии с коэффициентами, равными нулю.

Пример 26.5

Рассмотрим величины x и y^2 из примера 26.3. Там мы видели, что регрессия y^2 по x^2 линейна с коэффициентом ρ^2 и, следовательно, она нелинейна по x при $\rho \neq 0$. Однако поскольку $\mathbf{M}(x|y) = \rho y$, то

$$\mathbf{M}(x \mid y^2) = \frac{1}{2}\{\mathbf{M}(x \mid y) + \mathbf{M}(x \mid -y)\} = 0,$$

так что регрессия x по y^2 линейна с коэффициентом регрессии, равным нулю.

Метод наименьших квадратов и приближенная линейная регрессия

26.8 В примерах 26.3—26.5 одна или обе регрессии являлись точно линейными. Однако когда мы имеем дело не с теоретической совокупностью, а с наблюденной (и *a fortiori*, когда мы должны учитывать выборочные флуктуации), то очень редко встречается точная линейная регрессия. Тем не менее, как на рис. 26.1, регрессия может быть достаточно близкой к линейной. В последнем случае допустимо использовать линейную регрессию в качестве приближения. Таким образом, мы приходим к задаче «подгонки» прямой линии к регрессии y по x.

Когда отсутствуют особые соображения, выбор метода подгонки произволен, подобно тому как при описании данных произволен выбор между средним и медианой как мерами расположения. Если мы подгоняем регрессию y по x, то понятно, что отклонения точек (y, x) от подгоняемой прямой в некотором смысле должны быть малыми для того, чтобы эта прямая адекватно их описывала. Мы могли бы выбирать прямую, которая минимизировала бы сумму абсолютных отклонений точек от этой прямой, но тогда возникают обычные математические трудности, связанные со знаком модуля. Так же как раньше эти трудности

привели нас к отказу от среднего отклонения как меры разброса в пользу стандартного отклонения, так и теперь они приводят нас к предложению минимизировать сумму *квадратов* отклонений.

Осталось определить, какие отклонения следует брать: по оси y, по оси x или же «нормальные» отклонения, получаемые с помощью перпендикуляров, опускаемых из каждой точки на прямую. Поскольку мы рассматриваем зависимость y от x, то кажется естественным минимизировать сумму квадратов отклонений по оси y. Таким образом, мы вернулись к методу наименьших квадратов: мы выбираем «наилучшую» прямую регрессии y по x,

$$y = \alpha_2 + \beta_2 x, \tag{26.14}$$

так, чтобы была минимальной сумма квадратов отклонений n наблюдений от подгоняемой прямой регрессии, т. е. величина

$$S = \sum_{i=1}^{n} \{y_i - (\alpha_2 + \beta_2 x_i)\}^2. \tag{26.15}$$

Задача состоит в определении коэффициентов α_2 и β_2. В **19.4** мы уже рассматривали эту задачу в более общей форме. В использованных там матричных обозначениях (26.14) имеет вид

$$\underset{(n\times 1)}{\boldsymbol{y}} = \underset{(n\times 2)}{(\mathbf{1} \vdots \boldsymbol{x})} \underset{(2\times 1)}{\begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix}},$$

где $\mathbf{1}$ есть $(n \times 1)$-вектор, состоящий из единиц. Из (19.12) получаем решение

$$\begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix} = \{(\mathbf{1} \vdots \boldsymbol{x})'(\mathbf{1} \vdots \boldsymbol{x})\}^{-1} (\mathbf{1} \vdots \boldsymbol{x})' \boldsymbol{y} =$$

$$= \begin{pmatrix} n & \sum x \\ \sum x & \sum x^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum y \\ \sum xy \end{pmatrix} = \frac{1}{n\sum x^2 - (\sum x)^2} \begin{pmatrix} \sum x^2 \sum y - \sum x \sum xy \\ n\sum xy - \sum x \sum y \end{pmatrix}.$$

Таким образом,

$$\beta_2 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} = \frac{\mu_{11}}{\sigma_1^2},$$

так же как и в (26.11) для случая точной линейной регрессии. Далее,

$$\alpha_2 = \frac{\sum x^2 \sum y - \sum x \sum xy}{n\sum x^2 - (\sum x)^2} = \mu_y - \beta_2 \mu_x,$$

что эквивалентно (26.9). Итак, (26.14) представляется в виде

$$y - \mu_y = \beta_2 (x - \mu_x),$$

аналогичном (26.13).

Следовательно, мы пришли к заключению, что вычисление приближенной линии регрессии по методу наименьших квадратов дает те же результаты, что и в случае точной линейной регрессии.

Коэффициент корреляции

26.9 Учитывая результат из **26.8**, мы рассмотрим теперь общий случай, в котором регрессия не является точно линейной. Коэффициенты линейной регрессии (26.10) и (26.11) служат коэффициентами приближенных прямых регрессии, хотя иногда эти прямые могут оказаться точными.

Определим теперь по смешанному моменту *коэффициент корреляции* ρ, полагая

$$\rho = \mu_{11}/(\sigma_1 \sigma_2), \tag{26.16}$$

откуда, учитывая (26.10), (26.11) и (26.16), находим, что

$$\rho^2 = \beta_1 \beta_2. \tag{26.17}$$

Коэффициент ρ является симметричной функцией относительно x и y, каким и должен быть любой коэффициент взаимозависимости. Так как он представляет собой однородную функцию от центральных моментов, то он инвариантен относительно изменений начала координат и масштаба. Коэффициент ρ имеет тот же знак, что β_1 и β_2, поскольку у каждой из трех величин числитель равен μ_{11}, а все знаменатели положительны. Если $\mu_{11} = 0$, то $\rho = 0$. Из (26.17) видно, что $|\rho|$ есть среднее геометрическое между $|\beta_1|$ и $|\beta_2|$.

По неравенству Коши — Буняковского

$$\mu_{11}^2 = \left\{ \int\int (x - \mu_x)(y - \mu_y)\, dF \right\}^2 \leqslant$$

$$\leqslant \left\{ \int\int (x - \mu_x)^2\, dF \right\} \left\{ \int\int (y - \mu_y)^2\, dF \right\} = \sigma_1^2 \sigma_2^2,$$

так что

$$0 \leqslant \rho^2 \leqslant 1. \tag{26.18}$$

Равенство в правой части (26.18) выполняется (см. **2.7**) тогда и только тогда, когда $(x - \mu_x)$ и $(y - \mu_y)$ пропорциональны, т. е. x и y связаны строго линейной функциональной зависимостью. Следовательно, коэффициент ρ, по существу, равен ковариации μ_{11}, деленной на некоторую величину, гарантирующую, что ρ лежит в интервале $(-1, +1)$.

Легко показать, что угол между двумя линиями регрессии (26.12) и (26.13) есть

$$\theta = \operatorname{arctg}\left\{\frac{\sigma_1\sigma_2}{\sigma_1^2+\sigma_2^2}\left(\frac{1}{\rho}-\rho\right)\right\}. \qquad (26.19)$$

Поэтому, когда ρ изменяется от -1 до $+1$, θ возрастает от 0 до максимального значения $\pi/2$ при $\rho = 0$ и затем убывает вновь к 0. Следовательно, две линии регрессии совпадают ($\rho^2 = 1$) тогда и только тогда, когда x и y связаны строго линейной функциональной зависимостью. Угол между линиями регрессии является прямым тогда и только тогда, когда $\rho = 0$. В этом случае говорят, что x и y *некоррелированы*.

Можно показать, что

$$\rho^2 = \mathrm{D}(\alpha_2 + \beta_2 x)/\sigma_2^2 = \mathrm{D}(\alpha_1 + \beta_1 y)/\sigma_1^2, \qquad (26.20)$$

где D означает вычисленную дисперсию. Доказательство соотношения (26.20) предлагается читателю в качестве упражнения 26.13.

ρ как коэффициент взаимозависимости

26.10 Из **26.6** и примера 26.2 видно, что, в то время как независимость x и y влечет равенство $\mu_{11} = \rho = 0$, обратное утверждение, вообще говоря, неверно. Оно выполняется для совместно нормальных величин и в некоторых других случаях (пример 26.1). В этом и заключается трудность интерпретации ρ как коэффициента взаимозависимости в общем случае. Действительно, мы видели, что ρ является, по существу, коэффициентом *линейной* взаимозависимости и не может отражать более сложных форм взаимозависимости. В общей постановке задача взаимосвязи слишком сложна для того, чтобы ее можно было описать единственным коэффициентом.

Кроме того, *выразить* некоторое качество — это не то же самое, что *измерить* его. В случае, когда равенство $\rho = 0$ влечет независимость, мы знаем из **26.9**, что с возрастанием $|\rho|$ возрастает и взаимозависимость, достигая предельного случая линейной функциональной зависимости при $|\rho| = 1$. Но даже здесь остается открытым вопрос, какую функцию от ρ нужно использовать в качестве *меры* взаимозависимости: из (26.20) видно, что величину ρ^2 легче интерпретировать, чем ρ, поскольку она является отношением дисперсии подгоняемой прямой к полной дисперсии. Оставляя в стороне вопрос интерпретации, можно сказать, что в таких случаях величина ρ дает нам некоторую меру, хотя могут существовать и лучшие меры С другой стороны, если равенство $\rho = 0$ не влечет независимости, то трудно интерпретировать коэффициент ρ как меру взаимозависимости и, пожа-

луй, разумнее рассматривать его как индикатор, чем как точную меру. В практической работе мы рекомендовали бы использовать ρ в качестве *меры* взаимозависимости только в случае нормального или близкого к нему распределения.

Вычисление коэффициентов

26.11 Из (26.10), (26.11) и (26.16) видно, что для вычисления коэффициентов регрессии и корреляции требуется знать две дисперсии и ковариацию μ_{11}. В **2.19** и в примере 2.7 мы уже обсуждали, как вычислять дисперсии. Ковариация вычисляется аналогично, с использованием тождества, установленного в (26.3),

$$\mu_{11} = \sum_{i=1}^{n} (x_i - \mu_x)(y_i - \mu_y)/n \equiv \sum_{i=1}^{n} x_i y_i/n - \mu_x \mu_y \equiv$$

$$\equiv \sum_{i=1}^{n} x_i y_i/n - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)/n^2.$$

Ради удобства мы часто берем произвольные начала координат a, b для x и y соответственно. Тогда тождественно по a и b

$$\mu_{11} = \sum (x-a)(y-b)/n - \{\sum (x-a)\}\{\sum (y-b)\}/n^2. \quad (26.21)$$

Иными словами, ковариация μ_{11} инвариантна при изменении начала координат, так как она является смешанным центральным моментом. Соотношение (26.21) выполняется, если положить $(x-a) \equiv (y-b)$, при этом оно переходит в выражение (2.21) для вычисления дисперсий. Обычно бывает удобно брать произвольную единицу измерения u_x для x и другую произвольную единицу u_y для y. Легко видеть, что это преобразование равносильно делению μ_{11} на $u_x u_y$, σ_1^2 — на u_x^2 и σ_2^2 — на u_y^2. Таким образом, β_1 будет умножаться на u_y/u_x, β_2 — на u_x/u_y, а ρ не меняется при изменении масштабов.

Итак, μ_{11}, σ_1^2 и σ_2^2, а следовательно, и β_1, β_2 и ρ не меняются при изменениях начала координат. Если для вычислительных целей вводятся *различные* масштабные множители, то β_1 и β_2 должны умножаться на соответствующие отношения; если для обеих переменных используется *один и тот же* масштабный множитель, то β_1 и β_2 не изменяются. Коэффициент ρ не изменяется при любом изменении масштаба.

Пример 26.6. Вычисление коэффициентов для группированных данных

Для группированных данных, таких, как в таблице 26.1, примем ширину группы по каждой переменной за единицу для этой переменной (если по какой-то переменной группы имеют неодинаковые размеры, то возьмем из них *наименьший*). Выберем так-

же начало координат для каждой переменной где-нибудь возле среднего, оцененного на глаз. Так, мы примем за начало координат по x в таблице 26.1 точку 64, центр модальной частотной группы, относительно которого маргинальное распределение величины x близко к симметричному; начало координат для y поместим в 134,5, поскольку для такого скошенного маргинального распределения, как распределение величины y, среднее должно лежать заметно выше модальной частотной группы (122,5). Размеры групп (2 и 6) принимаем за единицы измерений. Сумма произведений $\sum xy$ вычисляется умножением каждой частоты по очереди на ее «координаты» в таблице, взятые в выбранных нами единицах. Например, частота 4, для которой $x = 68$, $y = 110{,}5$, умножается на $(+2)\cdot(-4) = -8$, так что ее вклад в сумму равен -32. Получаем

$$\sum x = -2{,}353, \qquad \sum y = -1{,}400,$$
$$\sum x^2 = 10{,}161, \qquad \sum y^2 = 70{,}802,$$
$$\sum xy = +8{,}786,$$

откуда

$$\mu_x = -\frac{2353}{4995}\cdot 2 + 64 = 63{,}06,$$

$$\mu_y = -\frac{1400}{4995}\cdot 6 + 134{,}5 = 132{,}82,$$

$$\sigma_1^2 = \left\{\frac{10{,}161}{4995} - \left(\frac{2353}{4995}\right)^2\right\}\times 2^2 = 7{,}25,$$

$$\sigma_2^2 = \left\{\frac{70\,802}{4995} - \left(\frac{1400}{4995}\right)^2\right\}\times 6^2 = 507{,}46,$$

$$\mu_{11} = \left\{\frac{8786}{4995} - \left(-\frac{2353}{4995}\right)\left(-\frac{1400}{4995}\right)\right\}\times 2\times 6 = +19{,}52,$$

и, следовательно,

$$\rho = \frac{\mu_{11}}{\sigma_1\sigma_2} = 0{,}322,$$
$$\beta_1 = \mu_{11}/\sigma_2^2 = 0{,}0385,$$
$$\beta_2 = \mu_{11}/\sigma_1^2 = 2{,}692.$$

Уравнения (приближенных) линейных регрессий имеют вид

$$x \text{ по } y\text{: } x - 63{,}06 = 0{,}0385\,(y - 132{,}82),$$

или

$$x = 0{,}0385y + 57{,}95,$$

$$y \text{ по } x\text{: } y - 132{,}82 = 2{,}692\,(x - 63{,}06),$$

или

$$y = 2{,}692x - 36{,}96.$$

Эти линии изображены на рис. 26.1 (стр. 378) как RR' и CC' соответственно.

Пример 26.7. Вычисление коэффициентов для негруппированных данных

В таблице 26.3 приведены данные о размерах урожаев пшеницы и картофеля по 48 графствам Англии в 1936 г. При негруппированных данных, таких, как эти, выбор некоторых новых начал координат и единиц измерений редко бывает удобным для

Таблица 26.3

Размеры урожаев пшеницы и картофеля по 48 графствам Англии в 1936 г.

Графство	Пшеница (центнеры *) на акр)	Картофель (тонны на акр)	Графство	Пшеница) (центнеры *) на акр)	Картофель (тонны на акр)
Бедфордшир	16,0	5,3	Нортгемптоншир	14,3	4,9
Хантингдоншир	16,0	6,6	Питерборо	14,4	5,6
Кембриджшир	16,4	6,1	Бакингемшир	15,2	6,4
Или	20,5	5,5	Оксфордшир	14,1	6,9
Суффолк Западный	18,2	6,9	Уорикшир	15,4	5,6
Суффолк Восточный	16,3	6,1	Шропшир	16,5	6,1
Эссекс	17,7	6,4	Вустершир	14,2	5,7
Хартфордшир	15,3	6,3	Глостершир	13,2	5,0
Мидлсекс	16,5	7,8	Уилтшир	13,8	6,5
Норфолк	16,9	8,3	Херефордшир	14,4	6,2
Линкс (Голландия)	21,8	5,7	Сомерсетшир	13,4	5,2
» (Кестевен)	15,5	6,2	Дорсетшир	11,2	6,6
» (Линдсей)	15,8	6,0	Девоншир	14,4	5,8
Йоркшир (Ист Райдинг)	16,1	6,1	Корнуолл	15,4	6,3
Кент	18,5	6,6	Нортамберленд	18,5	6,3
Суррей	12,7	4,8	Дарем	16,4	5,8
Суссекс Восточный	15,7	4,9	Йоркшир (Норт Райдинг)	17,0	5,9
Суссекс Западный	14,3	5,1	» (Уэст Райдинг)	16,9	6,5
Беркшир	13,8	5,5	Камберленд	17,5	5,8
Гемпшир	12,8	6,7	Уэстморленд	15,8	5,7
О-в Уайт	12,0	6,5	Ланкашир	19,2	7,2
Нотингемшир	15,6	5,2	Чешир	17,7	6,5
Лестершир	15,8	5,2	Дербишир	15,2	5,4
Ратленд	16,6	7,1	Стаффордшир	17,1	6,3

*) Английский центнер равен 50,8 *кг*, 1 акр = 0,405 *га*. (*прим. ред.*)

вычислений. Поэтому, используя обычные начала координат и единицы измерений, находим

$$\sum x = 758{,}0, \qquad \mu_x = 15{,}792, \qquad \beta_1 = \mu_{11}/\sigma_2^2 = 0{,}612,$$
$$\sum y = 291{,}1, \qquad \mu_y = 6{,}065, \qquad \beta_2 = \mu_{11}/\sigma_1^2 = 0{,}078,$$
$$\sum x^2 = 12170{,}48, \qquad \sigma_1^2 = 4{,}174, \qquad \rho = \mu_{11}/(\sigma_1\sigma_2) = 0{,}219.$$
$$\sum y^2 = 1791{,}03, \qquad \sigma_2^2 = 0{,}534,$$
$$\sum xy = 4612{,}64, \qquad \mu_{11} = 0{,}327.$$

Следовательно, уравнения линий регрессии (приближенных) имеют вид

Регрессия x по y: $x - 15{,}792 = 0{,}612(y - 6{,}065)$.
Регрессия y по x: $y - 6{,}065 = 0{,}078(x - 15{,}792)$.

Данные и прямые регрессии изображены на рис. 26.3, причем каждая точка на диаграмме соответствует паре (x, y). Диаграмма, подобная приведенной, на которую нанесены все точки,

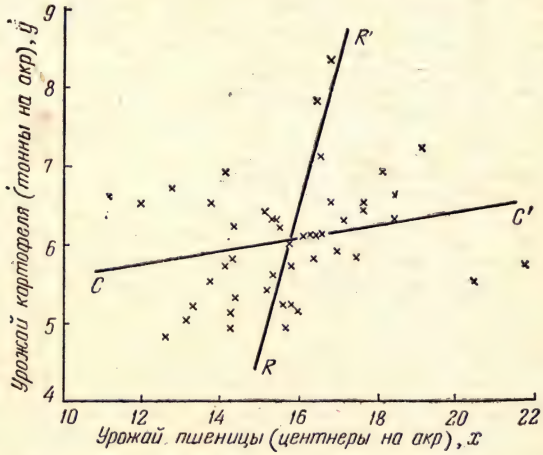


Рис. 26.3. Данные таблицы 26.3 с линиями регрессии.

называется *диаграммой разброса*; пользоваться ею мы очень рекомендуем, так как на ней можно быстро и просто увидеть, насколько хорошо соответствуют данным подгоняемые прямые регрессии (в нашем примере соответствие не очень хорошее). Действительно, диаграмма разброса, построенная заранее, до анализа, может дать ответ на вопрос, стоит ли вообще заниматься подгонкой линий регрессии.

Выборочные коэффициенты и их стандартные ошибки

26.12 Перейдем теперь к рассмотрению задач о выборочных коэффициентах корреляции и регрессии. Как обычно, будем соблюдать соглашение о том, что латинской буквой обозначается выборочная статистика, а греческой — ее эквивалент по генеральной совокупности. Таким образом, для выборочных коэффициентов регрессии и корреляции имеем

$$\left.\begin{aligned} b_1 &= m_{11}/s_2^2 = \frac{1}{n}\sum (x-\bar{x})(y-\bar{y}) \Big/ \frac{1}{n}\sum (y-\bar{y})^2, \\ b_2 &= m_{11}/s_1^2 = \frac{1}{n}\sum (x-\bar{x})(y-\bar{y}) \Big/ \frac{1}{n}\sum (x-\bar{x})^2, \\ r &= m_{11}/(s_1 s_2) = \frac{1}{n}\sum (x-\bar{x})(y-\bar{y}) \Big/ \left\{ \frac{1}{n}\sum (x-\bar{x})^2 \frac{1}{n}\sum (y-\bar{y})^2 \right\}^{1/2}, \end{aligned}\right\} \tag{26.22}$$

где суммирование происходит по всем выборочным значениям. Для вычислительных целей, так же как в случае генеральных коэффициентов β_1, β_2, ρ, эти выражения можно представить в виде

$$\left.\begin{aligned} b_1 &= \frac{\sum xy - (\sum x)(\sum y)/n}{\sum y^2 - (\sum y)^2/n}, \\ b_2 &= \frac{\sum xy - (\sum x)(\sum y)/n}{\sum x^2 - (\sum x)^2/n}, \\ r &= \frac{\sum xy - (\sum x)(\sum y)/n}{\left[\{\sum x^2 - (\sum x)^2/n\}\{\sum y^2 - (\sum y)^2/n\}\right]^{1/2}}. \end{aligned}\right\} \tag{26.23}$$

Как и раньше, $-1 \leqslant r \leqslant +1$.

26.13 Легко найти стандартные ошибки коэффициентов (26.22). Действительно, в примере 10.6 уже была получена асимптотическая дисперсия величины r. Там мы видели, что это (в общем случае) есть выражение, включающее в себя все вторые и четвертые моменты совокупности, из которой извлекается выборка. Однако в нормальном случае мы нашли, что оно сводится к

$$\mathrm{D}r = (1-\rho^2)^2/n. \tag{26.24}$$

Выражение (26.24) имеет для практики небольшую ценность, поскольку распределение коэффициента r слишком медленно сходится к нормальному (см. **16.29**): неразумно пользоваться нормальным приближением при $n < 500$. Эта трудность практи-

чески несущественна, так как (см. **16.33**) простое преобразование статистики r

$$z = \frac{1}{2}\log\left(\frac{1+r}{1-r}\right) = \operatorname{arg\,th} r \tag{26.25}$$

дает величину, распределение которой (для нормальной выборки) значительно ближе к нормальному с приближенным средним

$$\mathbf{M}(z) \approx \log\left(\frac{1+\rho}{1-\rho}\right) \tag{26.26}$$

и приближенной дисперсией

$$\mathbf{D}z \approx \frac{1}{n-3}, \tag{26.27}$$

не зависящей от ρ. Для $n > 50$ допустимо использование этой стандартной ошибки величины z. Более точные аппроксимации даны в **16.33**.

В случае выборочного коэффициента регрессии y по x

$$b_2 = m_{11}/s_1^2$$

использование (10.17) приводит, в точности как для r в примере 10.6, к выражению

$$\mathbf{D}b_2 = \left(\frac{\mu_{11}}{\sigma_1^2}\right)^2 \left\{ \frac{\mathbf{D}m_{11}}{\mu_{11}^2} + \frac{\mathbf{D}(s_1^2)}{\sigma_1^4} - \frac{2\operatorname{cov}(m_{11}, s_1^2)}{\mu_{11}\sigma_1^2} \right\}.$$

Подставляя сюда дисперсии и ковариацию из (10.23) и (10.24), находим

$$\mathbf{D}b_2 = \frac{1}{n}\left(\frac{\mu_{11}}{\sigma_1^2}\right)^2 \left\{ \frac{\mu_{22}}{\mu_{11}^2} + \frac{\mu_{40}}{\sigma_1^4} - \frac{2\mu_{31}}{\mu_{11}\sigma_1^2} \right\}. \tag{26.28}$$

Если исходная совокупность нормальна, то, учитывая соотношения из примера 3.17, получаем

$$\mathbf{D}b_2 = \frac{1}{n}\left(\frac{\mu_{11}}{\sigma_1^2}\right)^2 \left\{ \frac{(1+2\rho^2)\,\sigma_1^2\sigma_2^2}{\mu_{11}^2} + \frac{3\sigma_1^2}{\sigma_1^4} - \frac{6\rho\sigma_1^3\sigma_2}{\mu_{11}\sigma_1^2} \right\} =$$

$$= \frac{1}{n}\left(\frac{\mu_{11}}{\sigma_1^2}\right)^2 \left\{ \frac{\sigma_1^2\sigma_2^2}{\mu_{11}^2} - 1 \right\} = \frac{1}{n}\frac{\sigma_2^2}{\sigma_1^2}(1-\rho^2). \tag{26.29}$$

Аналогично, для коэффициента регрессии x по y имеем

$$\mathbf{D}b_1 = \frac{1}{n}\frac{\sigma_1^2}{\sigma_2^2}(1-\rho^2). \tag{26.30}$$

Выражения (26.29) и (26.30) (когда в них заменены σ_1^2, σ_2^2 и ρ на s_1^2, s_2^2 и r) более полезны для вычисления стандартных ошибок, чем (26.24), потому что, как мы видели в **16.35**, распределение величины b_2 симметрично относительно β_2. Читателю предлагается в качестве упражнения 26.9 показать с помощью (16.86), что соотношение (26.29) является точным, если его умножить на $n/(n-3)$, и что распределение статистики b_2 быстро сходится к нормальному ввиду того, что ее эксцесс есть величина порядка $1/n$.

Оценивание ρ в нормальных выборках

26.14 Выборочная теория двумерного нормального распределения изложена в **16.23—36**. Опираясь на наши результаты из глав 17 и 18, мы можем рассмотреть задачу оценивания коэффициента ρ по выборке.

В **16.24** было фактически показано, что функция правдоподобия дается выражением (16.52), в которое наблюдения входят только через пять статистик: $\bar{x}$, $\bar{y}$, s_1^2, s_2^2, r. Следовательно, они образуют систему достаточных статистик для пяти параметров: μ_1, μ_2, σ_1^2, σ_2^2, ρ, и согласно **23.10** их распределение полно. Далее, из (16.52) вытекает, что, даже если первые четыре параметра известны, тем не менее для одного ρ требуется эта пятикомпонентная достаточная статистика.

В главе 18 было показано, что оценка максимального правдоподобия коэффициента ρ принимает различные формы в зависимости от того, какие из остальных параметров оцениваются одновременно с ρ: МП-оценка всегда является функцией от системы достаточных статистик, но в разных ситуациях это *разные* функции. Когда оценивается только ρ, МП-оценка есть корень кубического уравнения (пример 18.3); когда оцениваются все пять параметров, то МП-оценка представляет собой выборочный коэффициент корреляции r (пример 18.14). На практике последний случай встречается наиболее часто, и мы теперь рассмотрим оценивание ρ с помощью r или некоторых функций от него.

26.15 Точное распределение коэффициента r, которое зависит только от ρ, дается в (16.61) или в более удобной форме в (16.66). Его среднее значение определяется гипергеометрическим рядом (16.73). Раскладывая в (16.73) гамма-функции в ряды Стирлинга (3.64) и беря два главных члена гипергеометрической функции, находим

$$\mathrm{M}(r) = \rho\left\{1 - \frac{(1-\rho^2)}{2n} + O(n^{-2})\right\}. \qquad (26.31)$$

Таким образом, r является несколько смещенной оценкой для ρ, когда $0 \neq \rho^2 \neq 1$. Смещение, как правило, мало, но интересно исследовать вопрос о том, можно ли его устранить.

26.16 Есть два подхода к этой задаче. Прежде всего, можно задать вопрос: существует ли функция $g(r)$ такая, что равенство

$$\mathbf{M}\{g(r)\} = g(\rho) \tag{26.32}$$

выполняется тождественно по ρ? Хотеллинг (1953) показал, что если g не зависит от n, то $g(r)$ может быть только линейной функцией от $\arcsin r$, а Харли (1956, 1957) доказал, что в самом деле

$$\mathbf{M}(\arcsin r) = \arcsin \rho. \tag{26.33}$$

Простое доказательство результата Харли дано Даниэлсом и Кендаллом (1958).

26.17 Второй, более прямой подход состоит в том, чтобы искать функцию от r, являющуюся несмещенной оценкой самого ρ. Согласно результату Хотеллинга из **26.16** эта функция должна зависеть от n. Так как r является функцией от системы полных достаточных статистик, то несмещенная функция от r должна быть единственной (см. **23.9**). Олкин и Прэтт (1958) нашли несмещенную оценку для ρ (обозначим ее r_u), представимую в виде гипергеометрической функции

$$r_u = rF\left[\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}(n-2), (1-r^2)\right]. \tag{26.34}$$

Раскладывая ее в ряд, получаем

$$r_u = r\left\{1 + \frac{1-r^2}{2(n-2)} + \frac{9(1-r^2)^2}{8n(n-2)} + O(n^{-3})\right\}. \tag{26.35}$$

Все члены ряда неотрицательны, поэтому

$$|r_u| \geqslant |r|,$$

и равенство выполняется только тогда, когда $r^2 = 0$ или 1. Тем не менее поскольку $F\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}(n-2), 0\right) = 1$ и r_u — возрастающая функция от r, то $r^2 \leqslant r_u^2 \leqslant 1$.

Очевидно, первый поправочный член в (26.35) нейтрализует отрицательное смещение порядка $1/n$ в (26.31). Олкин и Прэтт рекомендуют использовать в качестве оценки величину

$$r_u^* = r\left\{1 + \frac{1-r^2}{2(n-4)}\right\}. \tag{26.36}$$

Выражение в фигурных скобках (26.36) отличается от r_u/r не более чем на 0,01 при $n \geqslant 8$ и не более чем на 0,001 при $n \geqslant 18$ равномерно по r.

Олкин и Прэтт приводят точные таблицы r_u для $n = 2(2)30$ и $|r| = 0(0{,}1)1$, которые показывают, что если $n \geqslant 14$, то $|r_u|$ никогда не превосходит $|r|$ более чем на 5%.

В заключение отметим, что поскольку n входит только в знаменатель гипергеометрического ряда, то $r_u \to r$ при $n \to \infty$, так что эта статистика имеет такое же предельное распределение, как и r, а именно нормальное распределение со средним ρ и дисперсией $(1-\rho^2)^2/n$.

Доверительные границы и критерии для ρ

26.18 Если проверяется гипотеза о том, что $\rho = 0$, то критерии, основанные на r, являются РНМН (см. упражнение 31.21); но это не так, когда проверяется гипотеза о ненулевом значении ρ. Однако, как показал Леман (1959), если ограничиться статистиками критериев, которые инвариантны относительно преобразований сдвига и масштаба, то односторонние критерии, основанные на r, будут РНМ инвариантными критериями.

Для построения интервальных оценок можно использовать графики Ф. Дэвида (1938), описанные в **20.21**. Вследствие замеченной в **23.26** двойственности между доверительными интервалами и критериями, эти графики могут быть использованы для получения значений ρ, отклоняемых критерием размера α, т. е. всех значений ρ, не покрываемых доверительным интервалом, соответствующим данному α. Ф. Дэвид (1937) показал, что этот критерий несколько смещен (соответственно смещены и доверительные интервалы). Последнее легче всего понять с помощью критерия, основанного на z-преобразовании, приведенного в **26.13**: если бы он был точным и величина z имела бы *точное* нормальное распределение с дисперсией, не зависящей от ρ, то проверка гипотезы о ρ была бы равносильна проверке гипотезы о среднем значении нормального распределения с известной дисперсией.

Как известно из примера 23.11, критерий с «равными хвостами» для этой гипотезы является несмещенным. Поскольку z есть взаимно однозначная функция от r, то критерий с «равными хвостами» для r тоже был бы несмещенным. Таким образом, незначительное смещение r-критерия с «равными хвостами» можно рассматривать как отражение приближенного характера z-преобразования.

Упражнение 26.15 показывает, что критерий ОП основывается на величине r, но он имеет неодинаковые «хвосты», за исключением случая, когда проверяется гипотеза $\rho = 0$.

26.19 С другой стороны, можно построить приближенный критерий, используя z-преобразование Фишера, простейшие сведения о котором приведены в **26.13**: для проверки гипотетического значения ρ вычисляем величину (26.25) и проверяем гипотезу о том, что она распределена нормально со средним (26.26) и дисперсией (26.27). Критерий с одним или двумя «хвостами» вы-

бирается в зависимости от того, является ли альтернатива к этой простой гипотезе односторонней или двусторонней.

Тем же путем, используя z-преобразование, можно проверять сложную гипотезу о равенстве коэффициентов корреляции двух двумерных нормальных совокупностей, из которых взяты независимые выборки. Действительно, каждая из двух преобразованных статистик z_1, z_2 будет распределена, как в **26.13**, а величина $(z_1 - z_2)$ будет иметь среднее нуль и дисперсию $1/(n_1 - 3) + 1/(n_2 - 3)$, где n_1 и n_2 — объемы выборок. Упражнения 26.19—26.21 показывают, что величина $(z_1 - z_2)$ представляет собой точную статистику ОП, когда $n_1 = n_2$, и приближенную, когда $n_1 \neq n_2$. Однако в обоих случаях критерий оказывается приближенным, будучи критерием, основанным на стандартной ошибке.

Таким способом нельзя проверять более общую сложную гипотезу о том, что два коэффициента корреляции ρ_1, ρ_2, отличаются на константу Δ. Ибо тогда величина

$$\mathbf{M}(z_1 - z_2) = \frac{1}{2}\log\left(\frac{1+\rho_1}{1-\rho_1}\right) - \frac{1}{2}\log\left(\frac{1+\rho_2}{1-\rho_2}\right) = \\ = \frac{1}{2}\log\left\{\left(\frac{1+\rho_1}{1-\rho_1}\right)\left(\frac{1-\rho_2}{1+\rho_2}\right)\right\}$$

не есть функция только от $|\rho_1 - \rho_2|$. Можно было бы использовать z-преобразование для проверки гипотезы

$$H_0\colon \frac{1+\rho_1}{1-\rho_1} = a\left(\frac{1+\rho_2}{1-\rho_2}\right)$$

при любой константе a, но эта гипотеза неинтересна. Так как не исследовано точное распределение разности $r_1 - r_2$, то нет способов проверки гипотезы H_0: $\rho_1 - \rho_2 = \Delta$, исключая, разумеется, случай очень большой выборки, когда могут быть использованы стандартные ошибки.

Критерии независимости и регрессионные критерии

26.20 В частном случае, когда хотят проверить гипотезу $\rho = 0$, т. е. проверить, являются ли нормальные величины *независимыми*, можно использовать точный результат из **16.28** о том, что статистика

$$t = \{(n-2)\, r^2/(1-r^2)\}^{1/2} \qquad (26.37)$$

имеет t-распределение Стьюдента с $n - 2$ степенями свободы. Величина t^2, по существу, является статистикой критерия ОП для гипотезы $\rho = 0$ (см. упражнение 26.15), и это эквивалентно критерию с «равными хвостами», основанному на t.

По существу, мы здесь проверяем гипотезу о равенстве нулю μ_{11}, ковариации совокупности, что, очевидно, влечет обращение в нуль коэффициентов регрессии β_1, β_2. В **16.36** было показано, что величина

$$t = (b_2 - \beta_2)\left\{\frac{s_1^2(n-2)}{(s_2^2 - b_2^2 s_1^2)}\right\}^{1/2} \tag{26.38}$$

имеет распределение Стьюдента с $n-2$ степенями свободы. Таким образом, $\mathbf{M}(b_2) = \beta_2$. Когда $\beta_2 = 0$, то, как мы уже видели, (26.38) совпадает с (26.37). Итак, критерий независимости может рассматриваться как критерий равенства нулю коэффициента регрессии, являющийся частным случаем общего критерия (26.38) для гипотез, касающихся β_2. Заметим, что точный критерий о любом гипотетическом значении β_2 много проще, чем критерий для ρ.

В главе 31 мы увидим, что критерии независимости можно строить без каких-либо предположений о нормальности исходного распределения.

Корреляционные отношения и линейность регрессии

26.21 В **26.8** обсуждался вопрос подгонки прямых к линиям регрессии, когда последние не являются строго линейными. Можно получить дальнейшие результаты о линейности регрессии. Рассмотрим вначале случай строго линейной регрессии x по y, когда выполняется (26.12). Возводя (26.12) в квадрат и беря математические ожидания относительно y, получаем

$$\mathbf{D}\{\mathbf{M}(x \mid y)\} = \mathbf{M}_y\{(\mathbf{M}(x \mid y) - \mathbf{M}(x))^2\} = \beta_1^2\sigma_2^2 = \mu_{11}^2/\sigma_2^2. \tag{26.39}$$

Определим теперь *корреляционное отношение* x по y, η_1, положив

$$\eta_1^2 = \mathbf{D}\{\mathbf{M}(x \mid y)\}/\sigma_1^2, \tag{26.40}$$

что является отношением дисперсии средних значений x-сечений к дисперсии величины x. В отличие от ρ, функция η_1^2, очевидно, несимметрична относительно x и y. Из (26.40) следует, что величина η_1^2 инвариантна относительно перестановки x-сечений, поскольку $\mathbf{D}\{\mathbf{M}(x|y)\}$ не зависит от того, в каком порядке берутся значения $\mathbf{M}(x|y)$. В этом состоит ее существенное отличие от коэффициента ρ, который чувствителен к любому изменению порядка сечений.

Из (26.39) видно, что если регрессия строго линейна, то

$$\eta_1^2 = \mu_{11}^2/(\sigma_1^2\sigma_2^2) = \rho^2.$$

Рассмотрим теперь общий случай, в котором регрессия может не быть строго линейной. Имеем

$$\sigma_1^2 = \mathbf{M}[\{x - \mathbf{M}(x)\}]^2 = \mathbf{M}[(\{x - \mathbf{M}(x\,|\,y)\} + \{\mathbf{M}(x\,|\,y) - \mathbf{M}(x)\})^2] =$$
$$= \mathbf{M}[\{x - \mathbf{M}(x\,|\,y)\}^2] + \mathbf{M}[\{\mathbf{M}(x\,|\,y) - \mathbf{M}(x)\}^2] =$$
$$= \mathbf{M}[\{x - \mathbf{M}(x\,|\,y)\}^2] + \mathbf{D}\{\mathbf{M}(x\,|\,y)\}, \quad (26.41)$$

так как слагаемое с удвоенным произведением равно нулю:

$$2\mathbf{M}[\{x - \mathbf{M}(x\,|\,y)\}\{\mathbf{M}(x\,|\,y) - \mathbf{M}(x)\}] =$$
$$= \underset{y}{\mathbf{M}}[\{\mathbf{M}(x\,|\,y) - \mathbf{M}(x)\}\underset{x\,|\,y}{\mathbf{M}}\{x - \mathbf{M}(x\,|\,y)\}] = 0.$$

Таким образом, из (26.40) и (26.41) вытекает, что

$$0 \leqslant \eta_1^2 \leqslant 1 \quad (26.42)$$

и $\eta_1^2 = 1$ тогда и только тогда, когда $\mathbf{M}[\{x - \mathbf{M}(x\,|\,y)\}^2] = 0$, т. е. когда каждое наблюдение лежит на кривой регрессии и, следовательно, x и y связаны строго функционально. Далее,

$$\rho^2/\eta_1^2 = \mu_{11}^2/[\sigma_2^2 \mathbf{D}\{\mathbf{M}(x\,|\,y)\}]. \quad (26.43)$$

По неравенству Коши — Буняковского

$$\mu_{11}^2 = (\underset{x,\,y}{\mathbf{M}}[\{y - \mathbf{M}(y)\}\{x - \mathbf{M}(x)\}])^2 =$$
$$= (\underset{y}{\mathbf{M}}[\{y - \mathbf{M}(y)\}\{\mathbf{M}(x\,|\,y) - \mathbf{M}(x)\}])^2 \leqslant$$
$$\leqslant \underset{y}{\mathbf{M}}[\{y - \mathbf{M}(y)\}^2]\,\underset{y}{\mathbf{M}}[\{\mathbf{M}(x\,|\,y) - \mathbf{M}(x)\}^2] = \sigma_2^2 \mathbf{D}\{\mathbf{M}(x\,|\,y)\}, \quad (26.44)$$

и равенство выполняется тогда и только тогда, когда величина $\{y - \mathbf{M}(y)\}$ пропорциональна $\{\mathbf{M}(x\,|\,y) - \mathbf{M}(x)\}$, т. е. если $\mathbf{M}(x\,|\,y)$ является строго линейной функцией от y. Таким образом, согласно (26.43) и (26.44)

$$\rho^2/\eta_1^2 \leqslant 1, \quad (26.45)$$

где равенство выполняется только в случае строго линейной регрессии x по y. Следовательно, из (26.42), (26.45) и (26.18) получаем окончательно неравенства

$$0 \leqslant \rho^2 \leqslant \eta_1^2 \leqslant 1. \quad (26.46)$$

Можно следующим образом резюмировать наши результаты о неравенствах (26.46), приведенные в **26.9—10** и в этом пункте:

(а) $\rho^2 = 0$, если x и y независимы, но обратное неверно;

(б) $\rho^2 = \eta_1^2 = 1$ тогда и только тогда ,когда имеется строгая линейная функциональная зависимость x от y;

(в) $\rho^2 \leqslant \eta_1^2 = 1$ тогда и только тогда, когда имеется строгая нелинейная функциональная зависимость x от y;

(г) $\rho^2 = \eta_1^2 < 1$ тогда и только тогда, когда регрессия x по y строго линейна, но нет функциональной зависимости;

д) $\rho^2 < \eta_1^2 < 1$ указывает на то, что не существует функциональной зависимости и *некоторая* нелинейная кривая регрессии «подходит» лучше, чем «наилучшая» прямая линия, так как из (26.20) и (26.40) следует, что $\mathsf{D}\{\mathsf{M}(x|y)\} > \mathsf{D}(\alpha_1 + \beta_1 y)$, т. е. средние сечений более разбросаны, чем значения, даваемые наиболее подходящей линейной регрессией. (Конечно, может не существовать лучше подходящей *простой* регрессионной кривой.)

Поскольку величина η_1^2 не учитывает порядка x-сечений, то она не является мерой ни для какого конкретного типа зависимости x от y, но значение величины $\eta_1^2 - \rho^2$ служит индикатором нелинейности регрессии; важно подчеркнуть, что именно индикатором, а не мерой. Для того чтобы судить о ее значимости, нужно учитывать также и число сечений (вместе с числом наблюдений). Этот вопрос будет обсуждаться в **26.24.**

26.22 В случае регрессии y по x определим аналогично предыдущему корреляционное отношение

$$\eta_2^2 = \mathsf{D}\{\mathsf{M}(y \mid x)\}/\sigma_2^2, \tag{26.47}$$

где снова

$$0 \leqslant \rho^2 \leqslant \eta_2^2 \leqslant 1.$$

Так как $\eta_1^2 = 1$ тогда и только тогда, когда имеется строгая функциональная зависимость, то равенство $\eta_1^2 = 1$ влечет $\eta_2^2 = 1$, и наоборот *). Квадраты корреляционных отношений всегда превосходят ρ^2, но если регрессия x по y линейна, а y по x нелинейна, то мы будем иметь $\eta_1^2 = \rho^2 < \eta_2^2$, как это показано в следующем примере.

Пример 26.8

Рассмотрим снова ситуацию из примера 26.6. Регрессия x по y^2 там была линейной с коэффициентом регрессии 0, так что корреляция между x и y^2 тоже равна нулю. Мы нашли, что $\mathsf{M}(x|y^2) = 0$, откуда следует равенство $\mathsf{D}\{\mathsf{M}(x|y^2)\} = 0$. Поэтому корреляционное отношение x по y^2 равно 0, как и должно быть, поскольку коэффициент корреляции равен нулю и регрессия линейна.

*) Это неверно. Например, для распределения, в котором (x, y) принимает значения $(-3, -1)$, $(-1, -1)$, $(1, 1)$, $(3, 1)$ с вероятностями 1/4 каждое, имеем $\eta_1 = 4/5$, $\eta_2 = 1$. Дело в том, что из функциональной зависимости y от x следует функциональная зависимость x от y, только если функция $y(x)$ взаимно однозначна (например, строго монотонна). (*Прим. ред.*)

Регрессия же y^2 по x нелинейна: в примере 26.3 было найдено, что

$$\mathbf{M}(y^2 \mid x) = 1 + \rho^2(x^2 - 1).$$

Следовательно,

$$\mathbf{D}\{\mathbf{M}(y^2 \mid x)\} = \mathbf{M}[\{\rho^2(x^2 - 1)\}^2] = \rho^4 \mathbf{M}[\{x^2 - 1\}^2] = 2\rho^4$$

и $\sigma_2^2 = 2$. Таким образом, корреляционное отношение y^2 по x есть ρ^4, что всегда (когда $\rho \neq 0$) превосходит коэффициент корреляции между x и y^2, который равен нулю.

Если нужно вычислить корреляционные отношения по выборочным данным, то, подставляя выборочные дисперсии и дисперсии средних значений сечений в (26.40) и (26.41), получим выборочное корреляционное отношение x по y:

$$e_1^2 = \frac{\sum\limits_{i=1}^{k} n_i (\bar{x}_i - \bar{x})^2}{\sum\limits_{i=1}^{k} \sum\limits_{j=1}^{n_i} (x_{ij} - \bar{x})^2} \equiv \frac{\sum\limits_{i} n_i \bar{x}_i^2 - n\bar{x}^2}{\sum\limits_{i} \sum\limits_{j} x_{ij}^2 - n\bar{x}^2}, \tag{26.48}$$

где $\bar{x}_i$ — среднее значение i-го x-сечения, n_i — число наблюдений в этом сечении, k — количество сечений. Аналогичное выражение имеет место для e_2^2, выборочного корреляционного отношения y по x. Как и в случае генеральных коэффициентов, имеем

$$0 \leqslant r^2 \leqslant e_i^2 \leqslant 1, \quad i = 1, 2. \tag{26.49}$$

Пример 26.9. Вычисление корреляционного отношения

Вычислим корреляционное отношение y по x для данных таблицы 26.1, которые теперь будем рассматривать как выборку. Вычисления приведены в таблице 26.4.

В примере 26.6 было найдено, что среднее значение y равно $\bar{y} = 132{,}82$, а дисперсия y равна 507,46. Таким образом, из (26.48) находим корреляционное отношение y по x:

$$e_2^2 = \frac{88\,385\,915{,}97 - 4995\,(132{,}82)^2}{4995 \times 507{,}46} =$$

$$= \frac{88\,385\,915{,}97 - 88\,117\,544{,}25}{2\,534\,762{,}70} = \frac{268\,359{,}73}{2534\,762{,}70} = 0{,}106.$$

Это значение ненамного превосходит квадрат коэффициента

Таблица 26.4

Рост (x)	Средний вес в сечении ($\bar{y}_i$)	$\bar{y}_i^2$	n_i	$n_i\bar{y}_i^2$
54	92,50	8556,25	5	42781,25
56	111,41	12412,19	33	409602,27
58	122,05	14896,20	254	3783634,80
60	124,43	15482,82	813	12587532,66
62	130,22	16957,25	1340	22722715,00
64	134,59	18114,47	1454	26338439,38
66	140,48	19734,63	750	14800972,50
68	146,37	21424,18	275	5891649,50
70	157,32	24749,58	56	1385976,48
72	163,41	26702,83	11	293731,13
74	179,50	32220,25	4	128881,00
			$n = 4995$	88385915,97

корреляции

$$r^2 = (0{,}322)^2 = 0{,}104.$$

Рисунок 26.1 показывает, что линейная аппроксимация регрессии CC' действительно хорошая.

Проверка гипотез о корреляционных отношениях и о линейности регрессии

26.23 В **26.21** мы видели, что равенство $\eta_1^2 = \rho^2$ указывает на то, что для регрессии нельзя найти лучшей кривой, чем прямая линия, а следовательно, положительное значение величины $\eta_1^2 - \rho^2$ является индикатором нелинейности регрессии. Теперь, когда определены выборочные корреляционные отношения e_i^2, естественно задать вопрос: можно ли с помощью статистики $(e_1^2 - r^2)$ построить критерий для проверки линейности регрессии x по y? При этом мы получим также критерий для проверки гипотезы о том, что $\eta_1^2 = 0$, и установим связь между этими критериями и критерием для гипотезы $\rho = 0$, приведенным в (26.37). Эти задачи впервые были решены Р. А. Фишером.

Согласно (26.49) в правой части тождества

$$ns_1^2 \equiv ns_1^2 r^2 + ns_1^2(e_1^2 - r^2) + ns_1^2(1 - e_1^2) \qquad (26.50)$$

все члены положительны. Так как

$$\sum_{i=1}^{k}\sum_{l=1}^{n_i}\{(\bar{x}_i - \bar{x}) - b_1(\bar{y}_i - \bar{y})\}^2 \equiv \sum_i\sum_l (\bar{x}_i - \bar{x})^2 - r^2\sum_i\sum_l (x_{il} - \bar{x})^2,$$

то (26.50) можно переписать в виде

$$\sum_i \sum_j (x_{ij} - \bar{x})^2 \equiv r^2 \sum_i \sum_j (x_{ij} - \bar{x})^2 + \\ + \sum_i \sum_j \{(\bar{x}_i - \bar{x}) - b_1(\bar{y}_i - \bar{y})\}^2 + \sum_i \sum_j (x_{ij} - \bar{x}_i)^2. \quad (26.51)$$

Выражение (26.51) представляет собой разложение квадратичной формы от x_{ij} на три других таких же формы. Предположим теперь, что y_i фиксированы, а все величины x_{ij} не зависят одна от другой и распределены нормально с одинаковыми дисперсиями (без потери общности будем считать их равными единице). Вопрос о средних значениях x_{ij} временно оставим открытым.

Если выполняется гипотеза H_0 о том, что все x_{ij} имеют одинаковые средние, т. е. что кривая регрессии представляет собой прямую, параллельную оси y, то левая часть (26.51), как известно, подчиняется хи-квадрат распределению с $n - 1$ степенями свободы. Доказательство того, что ранги квадратичных форм в правой части (26.51) равны соответственно 1, $(k - 2)$ и $(n - k)$, проводится непосредственно, хотя оно и громоздко. Так как эти ранги в сумме дают $(n - 1)$, то из теоремы Кокрэна (**15.16**) вытекает, что три слагаемых справа независимы и имеют хи-квадрат распределения с такими же числами степеней свободы, каковы их ранги. Согласно **16.15** отношение любых двух из них (деленных на их ранги) подчиняется F-распределению с соответствующими степенями свободы. Этот факт может быть использован двояко для проверки гипотезы H_0.

(а) Отношение первого слагаемого, деленного на его ранг, к сумме второго и третьего, тоже деленной на ее ранг,

$$\frac{r^2/1}{(1 - r^2)/(n - 2)}, \quad \text{есть} \quad F_{1,\, n-2}, \quad (26.52)$$

где индексы указывают число степеней свободы.

Как можно заметить, это эквивалентно критерию (26.37), поскольку $t^2_{n-2} \equiv F_{1,\, n-2}$ согласно **16.15**. Это утверждение было получено в **16.28** для двумерного нормального распределения. Здесь мы считаем игреки фиксированными, а распределения внутри каждого x-сечения нормальными.

(б) Отношение суммы первого и второго слагаемых, деленной на ее ранг, к третьему, тоже деленному на его ранг,

$$\frac{e_1^2/(k - 1)}{(1 - e_1^2)/(n - k)}, \quad \text{есть} \quad F_{k-1,\, n-k}. \quad (26.53)$$

Для обоих критериев большие значения статистики критерия ведут к отклонению H_0.

Критерии, основанные на (26.52) и (26.53), совершенно различны, и оба имеют силу для H_0, но (26.52), по существу, про-

веряет гипотезу $\rho^2 = 0$, тогда как (26.53) — гипотезу $\eta_1^2 = 0$. Если альтернативная гипотеза состоит в том, что регрессия x по y линейна, то критерий (26.52) будет иметь бóльшую мощность. Если же альтернативная гипотеза утверждает, что регрессия может быть любой формы, отличной от указанной в H_0, то (26.53), очевидно, является более мощным. На практике почти всегда используется (26.52) в форме линейного регрессионного критерия (**26.20**), но существуют определенные ситуации, в которых более применим критерий, основанный на (26.53). Эти критерии мы обсудим ниже, в **26.24**, но сначала рассмотрим критерий линейности регрессии.

26.24 Если не все x_{ij} имеют одинаковые средние, то левая часть в (26.51) уже не подчиняется распределению χ^2_{n-1}. Однако если мы перенесем первое слагаемое справа в левую часть, то получим

$$ns_1^2(1 - r^2) \equiv ns_1^2(e_1^2 - r^2) + ns_1^2(1 - e_1^2). \qquad (26.54)$$

Поскольку

$$ns_1^2(1 - r^2) \equiv \sum_i \sum_j \{x_{ij} - (a_1 + b_1 y_i)\}^2$$

есть сумма квадратов уклонений от «подогнанной» линейной регрессии, то при гипотезе H_0' о том, что регрессия x по y строго линейна, и при нормальных, как и прежде, распределениях внутри сечений величина $ns_1^2(1 - r^2)$ имеет хи-квадрат распределение с $n - 2$ степенями свободы (при определении каждого из параметров линии регрессии теряется по одной степени свободы (см. **19.9**)). Ранги квадратичных форм в правой части (26.54) равны по-прежнему $(k - 2)$ и $(n - k)$. Следовательно, эти формы независимы и подчинены хи-квадрат распределениям с теми же числами степеней свободы. Таким образом, их отношение, после деления на соответствующие ранги,

$$\frac{(e_1^2 - r^2)/(k - 2)}{(1 - e_1^2)/(n - k)}, \quad \text{есть} \quad F_{k-2,\, n-k}. \qquad (26.55)$$

Выражение (26.55) может быть использовано для проверки гипотезы H_0' о линейности регрессии; H_0' отвергается при больших значениях статистики критерия. Снова мы не делали никаких предположений о величинах x_{ij}.

Итак, наши интуитивные соображения о том, что с помощью $(e_1^2 - r^2)$ можно проверять линейность регрессии, оказались верными. Но (26.55) показывает, что статистика критерия зависит от $(1 - e_1^2)$, k и n, так что одной величины $(e_1^2 - r^2)$ недостаточно.

Все три критерия, которые мы обсуждали здесь и в предыдущем пункте, являются критериями ОП линейных гипотез

вида, рассмотренного во второй части главы 24. Например, гипотеза о том, что все величины x_{ij} имеют одинаковые средние, допускает две интерпретации. Можно считать, что средние лежат на прямой линии, определяемой двумя параметрами, и проверять гипотезу о том, что эта линия имеет нулевой наклон. Последняя гипотеза содержит одно ограничение на два параметра. В обозначениях пунктов **24.27—28** $k = 2$ и $r = 1$, так что мы получаем F-критерий с $(1, n-2)$ степенями свободы: это критерий (26.52). С другой стороны, можно считать, что средние значения k сечений лежат на k-параметрической кривой (например, на кривой, определяемой полиномом степени $k-1$), и проверять гипотезу о том, что все коэффициенты полинома, кроме константы, равны нулю (гипотеза содержит $k-1$ ограничений). Тогда мы получаем F-критерий с $(k-1, n-k)$ степенями свободы: это критерий (26.53). Наконец, если во второй формулировке проверять гипотезу о том, что все коэффициенты полинома, кроме константы и коэффициента при линейном члене, равны нулю (так что средние значения сечений лежат на прямой линии), то мы наложим $k-2$ ограничений и получим F-критерий с $(k-2, n-k)$ степенями свободы: это уже критерий (26.55).

Следовательно, для фиксированных значений y_i результаты главы 24, касающиеся мощности критерия ОП, основанного на нецентральном F-распределении, применимы и к этим критериям, которые согласно **24.37** являются РНМ инвариантными критериями. Однако в двумерном нормальном случае, где y_i не фиксированы, получающиеся распределения *не* будут совпадать с распределениями для фиксированных, как выше, y_i, за исключением случая справедливости проверяемой гипотезы, когда изменение y_i является несущественным (как будет показано в **27.29**). Например, распределение величины r^2, полученное из нецентрального F-распределения для (26.52), не совпадает с результатом, получаемым из (16.61) или (16.66) в случае двумерной нормальности. Следовательно, функции мощности критерия для $\rho = 0$ в этих двух случаях различны, хотя один и тот же критерий применим в обоих случаях. Однако для больших n результаты совпадают. В **27.29** и **27.31** мы обсудим это в более общей ситуации в связи с множественным коэффициентом корреляции (r^2 есть частный случай последнего).

Внутриклассовая корреляция

26.25 Иногда возникают задачи, главным образом в биологических исследованиях, в которых требуется найти корреляцию между членами одного или большего числа семейств. Пусть, например, исследуется корреляция между ростами

братьев. Тогда перед нами встает вопрос: какая переменная является первой и какая — второй? В простейшем случае мы можем располагать некоторым числом семейств, каждое из которых содержит двух братьев. Наша корреляционная таблица имеет две переменные, обе обозначают рост, но при ее составлении мы должны решить, какой брат относится к какой переменной. Один из способов заключается в том, чтобы с первой переменной связывать либо старшего брата, либо более высокого; но этот способ дал бы нам корреляцию между старшими и младшими братьями или между более высокими и теми, что ниже, а не корреляцию между братьями вообще, которая нам требуется.

Решение задачи состоит в том, чтобы помещать в таблицу обе возможные пары, т. е. пары, получаемые, если каждого из братьев считать первым. Если семейство или, в более общем случае, класс содержит k членов, то в таблице будет $k(k-1)$ записей, так как каждый член будет браться первым совместно с каждым отличным от него членом, который берется вторым. При p классах с $k_1, k_2, \ldots, k_p$ членами корреляционная таблица будет иметь $\sum_{i=1}^{p} k_i(k_i-1) = N$ записей.

Таблица 26.5

Рост (дюймы)	Рост (дюймы)						
	68	69	70	71	72	73	Всего
68	—	—	2	—	—	—	2
69	—	—	1	—	1	—	2
70	2	1	2	1	2	—	8
71	—	—	1	—	4	1	6
72	—	1	2	4	2	1	10
73	—	—	—	1	1	—	2
Всего	2	2	8	6	10	2	30

В качестве простой иллюстрации рассмотрим пять семейств, содержащих по три брата, ростá которых (в дюймах) есть соответственно: 69, 70, 72; 70, 71, 72; 71, 72, 72; 68, 70, 70; 71, 72, 73. В таблице будет 30 записей (см. таблицу 26.5).

Здесь, к примеру, пара 69, 70 из первого семейства вошла в таблицу как (69, 70) и (70, 69), а пара 72, 72 из третьего семейства — дважды как (72, 72).

Очевидно, что таблица симметрична относительно главной диагонали, как и должно быть. Обычным способом можно вычислить коэффициент корреляции. Находим $\sigma_1^2=\sigma_2^2=1{,}716$, $\mu_{11}=0{,}516$ и, следовательно, $\rho=0{,}516/1{,}716=0{,}301$.

Такой коэффициент корреляции называется *внутриклассовым* коэффициентом корреляции. Его можно находить следующим более прямым способом.

Предположим, что имеется p классов со значениями $x_{11}, \ldots, x_{1k_1}; x_{21}, \ldots, x_{2k_2}; \ldots; x_{p1}, \ldots, x_{pk_p}$. В корреляционной таблице каждый член i-го класса будет появляться $k_i - 1$ раз (вместе с каждым другим элементом его класса). Таким образом, среднее значение каждой переменной равно

$$\mu=\frac{1}{N}\sum_{i=1}^{p}(k_i-1)\sum_{j=1}^{k_i}x_{ij},$$

а дисперсия

$$\sigma^2=\frac{1}{N}\sum_{i=1}^{p}(k_i-1)\sum_{j=1}^{k_i}(x_{ij}-\mu)^2.$$

Соответственно ковариация равна

$$\mu_{11}=\frac{1}{N}\sum_{i=1}^{p}\sum_{\substack{j,\,l=1\\ j\neq l}}^{k_i}(x_{ij}-\mu)(x_{il}-\mu)=$$

$$=\frac{1}{N}\left\{\sum_{i=1}^{p}\sum_{j,\,l=1}^{k_i}(x_{ij}-\mu)(x_{il}-\mu)-\sum_{i=1}^{p}\sum_{j=1}^{k_i}(x_{ij}-\mu)^2\right\}=$$

$$=\frac{1}{N}\left\{\sum_{i=1}^{p}\left[\sum_{j=1}^{k_i}(x_{ij}-\mu)\right]^2-\sum_i\sum_j(x_{ij}-\mu)^2\right\}=$$

$$=\frac{1}{N}\left\{\sum_i k_i^2(\mu_i-\mu)^2-\sum_i\sum_j(x_{ij}-\mu)^2\right\},$$

где μ_i — среднее значение i-го класса. Следовательно, для коэффициента корреляции получаем выражение

$$\rho = \frac{\sum_i k_i^2 (\mu_i - \mu)^2 - \sum_i \sum_j (x_{ij} - \mu)^2}{\sum_i (k_i - 1) \sum_j (x_{ij} - \mu)^2}. \qquad (26.56)$$

Если $k_i = k$ для всех i, то (26.56) преобразуется к виду

$$\rho = \frac{k^2 p \sigma_{\mu_i}^2 - kp\sigma^2}{(k-1)kp\sigma^2} = \frac{1}{k-1}\left(\frac{k\sigma_{\mu_i}^2}{\sigma^2} - 1\right), \qquad (26.57)$$

где $\sigma_{\mu_i}^2$ — дисперсия средних значений по классам, равная $\frac{1}{p}\sum_{i=1}^{p}(\mu_i - \mu)^2$.

Для того чтобы отличать внутриклассовый коэффициент от обычного коэффициента корреляции ρ, будем обозначать его ρ_i, а соответствующее выборочное значение r_i.

Пример 26.10

Используя (26.57), найдем внутриклассовый коэффициент для данных таблицы 26.5. Выберем начало координат в точке 70 дюймов. Тогда значения переменных равны —1, 0, 2; 0, 1, 2; 1, 2, 2; —2, 0, 0; 1, 2, 3. Следовательно, $\mu = 13/15$, $\mu_2' = 1/15\{(-1)^2 + 0^2 + \ldots\} = 37/15$ и $\sigma^2 = 386/225$. Средние значения по семействам, μ_i, равны

$$\frac{5}{15}, \quad \frac{15}{15}, \quad \frac{25}{15}, \quad \frac{-10}{15}, \quad \frac{30}{15},$$

а их отклонения от μ равны

$$\frac{-8}{15}, \quad \frac{2}{15}, \quad \frac{12}{15}, \quad \frac{-23}{15}, \quad \frac{17}{15}.$$

Таким образом,

$$\sigma_{\mu_i}^2 = \frac{1}{5}\left\{\left(\frac{-8}{15}\right)^2 + \ldots\right\} = \frac{1030}{1125}.$$

Следовательно, согласно (26.57)

$$\rho_i = \frac{1}{2}\left\{\frac{3 \cdot 1030 \cdot 225}{1125 \cdot 386} - 1\right\} = 0{,}301.$$

Этот результат был уже получен непосредственно в **26.25**.

26.26 При интерпретации внутриклассового коэффициента корреляции необходима осторожность. Из (26.57) видно, что ρ_i не может быть меньше, чем $\frac{-1}{k-1}$, однако он может достигать

значения $+1$, когда $\sigma^2_{\mu_i} = \sigma$. Таким образом, он является асимметричным коэффициентом в том смысле, что его отрицательное значение как мера отклонения от независимости не равнозначно соответствующему положительному значению.

Фактически внутриклассовый коэффициент со многих точек зрения удобнее рассматривать как величину, связанную простым линейным преобразованием с отношением дисперсий между классами и внутри классов в дисперсионном анализе. При таком подходе для случая, когда все семейства имеют одинаковые объемы k, Фишер (1921c) вывел распределение выборочного внутриклассового коэффициента r_i. Он нашел, что если $k = 2$, то, как и в случае коэффициента r, преобразование

$$z = \operatorname{arg\,th} r_i$$

дает статистику (z), распределение которой близко к нормальному со средним $\zeta = \operatorname{arg\,th} \rho_i$ и дисперсией, не зависящей от ρ_i. Для $k > 2$ необходимо более сложное преобразование. Результаты Фишера приведены в упражнении 26.14.

Тетрахорическая корреляция

26.27 Рассмотрим теперь оценивание коэффициента ρ в двумерной нормальной совокупности, когда данные известны не во всех подробностях. Возьмем прежде всего крайний случай, иллюстрируемый таблицей 26.6. Она основана на распределении коров по возрасту и удойности, приведенном в таблице 1.24 упражнения 1.6. Предположим, что вместо той таблицы нам была бы дана только

Таблица 26.6

Распределение коров по возрасту и удойности

	Возраст 6 и более лет	Возраст 3—5 лет	Всего
Удойность 8—18 галлонов *)	1078	1407	2485
Удойность 19 и более галлонов	1546	881	2427
Всего . . .	2624	2288	4912

*) 1 галлон = 4,54 *л.* (*Прим. ред.*)

Это крайне сокращенный вариант первоначальной таблицы. Предположим, что исходное распределение является двумерным нормальным. Как оценить ρ из этой таблицы? В общем

случае задача состоит в оценке ρ для таблицы 2×2 с частотами

$$\begin{array}{cc|c} a & b & a+b \\ c & d & c+d \\ \hline a+c & b+d & a+b+c+d=n \end{array} \tag{26.58}$$

В (26.58) мы всегда будем в качестве d брать такую частоту, чтобы никакая из отвечающих ей маргинальных частот не содержала медианного значения варианты.

Если эта таблица получена с помощью двойной дихотомии двумерного нормального распределения

$$f(x, y) \propto z_0 \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2} - \frac{2\rho xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)\right\}, \tag{26.59}$$

то можно найти такое h', что

$$\int_{-\infty}^{h'} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = \frac{a+c}{n}.$$

Полагая $h = h'/\sigma_1$, имеем

$$(2\pi)^{-1/2} \int_{-\infty}^{h} \exp\left(-\frac{1}{2}x^2\right) dx = \frac{a+c}{n}, \tag{26.60}$$

и, следовательно, h можно определить из таблиц одномерной нормальной функции распределения. Аналогично, существует такое k, что

$$(2\pi)^{-1/2} \int_{-\infty}^{k} \exp\left(-\frac{1}{2}y^2\right) dy = \frac{a+b}{n}. \tag{26.61}$$

По нашему соглашению о расположении данных в таблице (26.58) h и k никогда не бывают отрицательными.

Подогнав таким способом одномерные нормальные распределения к маргинальным частотам таблицы, мы должны теперь для нахождения ρ решить уравнение

$$\frac{d}{n} = \int_{h}^{\infty} \int_{k}^{\infty} z_0 \exp\left\{\frac{-1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right\} dx\, dy. \tag{26.62}$$

Подинтегральное выражение в (26.62) является нормированным, потому что h и k выбирались как нормированные отклонения. Разложим подинтегральную функцию по степеням ρ. Характеристическая функция распределения имеет вид

$$\varphi(t, u) = \exp\left\{-\frac{1}{2}(t^2 + 2\rho tu + u^2)\right\}.$$

Таким образом, используя двумерную форму теоремы обращения (**4.17**), находим из (26.62)

$$\frac{d}{n}=\int_h^\infty\int_k^\infty\left\{\frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(t,\,u)\exp(-itx-iuy)\,dt\,du\right\}dx\,dy=$$

$$=\int_h^\infty\int_k^\infty\left\{\frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}(t^2+u^2)-itx-iuy\right)\times\right.$$

$$\left.\times\sum_{j=0}^{\infty}\frac{(-\rho)^j t^j u^j}{j!}\,dt\,du\right\}dx\,dy. \quad (26.63)$$

Коэффициент при $(-\rho)^j/j!$ представляет собой произведение двух интегралов, из которых первый есть

$$I(x,\,h,\,t)=\int_h^\infty\left\{\frac{1}{2\pi}\int_{-\infty}^{\infty}t^j\exp\left(-\frac{1}{2}t^2-itx\right)dt\right\}dx, \quad (26.64)$$

а второй есть $I(y, k, u)$. Согласно **6.18** интеграл в фигурных скобках в (26.64) равен

$$(-i)^j H_j(x)\,\alpha(x),$$

где

$$\alpha(x)=(2\pi)^{-1/2}\exp\left(-\frac{1}{2}x^2\right).$$

Согласно (6.21)

$$-\frac{d}{dx}\{H_{j-1}(x)\,\alpha(x)\}=H_j(x)\,\alpha(x).$$

Следовательно, двойной интеграл в (26.64) имеет вид

$$I(x,\,h,\,t)=\left[(-1)^{j-1}i^j H_{j-1}(x)\,\alpha(x)\right]_h^\infty=(-i)^j H_{j-1}(h)\,\alpha(h). \quad (26.65)$$

Подставляя в (26.63) из (26.65) выражения для $I(x, h, t)$, $I(y, k, u)$, получаем ряд

$$\frac{d}{n}=\sum_{j=0}^{\infty}\frac{\rho^j}{j!}H_{j-1}(h)\,H_{j-1}(k)\,\alpha(h)\,\alpha(k). \quad (26.66)$$

Переписывая его в терминах тетрахорических функций, определенных в (6.44), находим

$$\frac{d}{n}=\sum_{j=0}^{\infty}\rho^j\tau_j(h)\,\tau_j(k). \quad (26.67)$$

26.28 Формально (26.67) представляет собой разрешимое относительно ρ уравнение, но на практике его решение с помощью последовательных приближений может оказаться очень громоздким. (Ряд (26.67) всегда сходится, хотя иногда и медленно.) Проще находить ρ интерполяцией по таблицам, содержащим значение интеграла d/n в зависимости от ρ при различных значениях h и k (*Tables for Statisticians and Biometricians*, Vol. 2).

Оценку для ρ, полученную таким путем по выборке объема n, называют *тетрахорическим* коэффициентом r. Мы будем обозначать ее r_t.

Пример 26.11

Для данных таблицы 26.6 находим, что нормальное отклонение, соответствующее $2624/4912 = 0{,}5342$, есть $h = 0{,}086$, и, аналогично, для $2485/4912 = 0{,}5059$ находим $k = 0{,}015$. Имеем также для d/n значение $881/4912 = 0{,}1794$.

С помощью таблиц при разных значениях h, k и ρ находим следующие значения величины d:

$\rho = -0{,}30$			$\rho = -0{,}35$		
	$h=0$	$h=0{,}1$		$h=0$	$h=0{,}1$
$k=0$	0,2015	0,1818	$k=0$	0,1931	0,1735
$k=0{,}1$	0,1818	0,1639	$k=0{,}1$	0,1735	0,1555

Для $h = 0{,}086$, $k = 0{,}015$ с помощью линейной интерполяции получаем приближенно $\rho = -0{,}32$. В таблице мы поменяли местами столбцы; учитывая это, имеем $\rho = +0{,}32$. Следовательно, мы получили $r_t = +0{,}32$. (Для таблицы 1.24 коэффициент корреляции $r = 0{,}22$.)

26.29 Тетрахорический коэффициент r_t используется главным образом психологами, у которых данные часто представляют собой таблицы 2×2. Для тетрахорического коэффициента в точной форме неизвестны выборочное распределение и даже его стандартная ошибка. Но Карл Пирсон (1913) нашел асимптотическое выражение стандартной ошибки. Существуют, однако, более простые способы вычислений, основанные на номограммах (Хейес (1946), Гамильтон (1948), Дженкинс (1955)) и таблицах для стандартной ошибки в приближенной форме (Гилфорд и Лайонс (1942), Хейес (1943); Гохеен и Каврук (1948)). По-видимому неизвестно, при каких объемах выборок

можно безбоязненно пользоваться такими стандартными ошибками.

Обобщение на полихорическое оценивание в $r \times c$ таблицах см. ниже, в **33.35**.

Бисериальная корреляция

26.30 Предположим теперь, что у нас есть таблица $2 \times q$, являющаяся дихотомией по некоторому качественному фактору и классификацией либо по числовой величине, либо также по качественному признаку, который может быть как упорядоченным, так и неупорядоченным.

Таблица 26.7 иллюстрирует рассматриваемый тип материала. В ней классифицированы 1426 преступников по тому, были они алкоголиками или нет, и по характеру преступления, за которое они были заключены в тюрьму.

Таблица 26.7

1426 преступников, классифицированных по типу преступления и отношению к алкоголизму.

(Данные Горинга, цитированные К. Пирсоном, 1909)

	Поджог	Изнасилование	Насильственные действия	Воровство	Изготовление фальшивых денег	Мошенничество	Всего
Алкоголики	50	88	155	379	18	63	753
Неалкоголики	43	62	110	300	14	144	673
Всего	93	150	265	679	32	207	1426

Хотя столбцы таблицы находятся в произвольном порядке (у нас они упорядочены по связи преступления с интеллектом, но этот порядок достаточно относителен), мы можем, однако, получить оценку для ρ, предполагая, что подлежащее исследованию двумерное распределение нормально. Действительно, при таком распределении $\rho^2 = \eta^2$, обе регрессии линейны и, как отмечено в **26.21**, статистика η^2 инвариантна относительно перестановки сечений. Мы переходим, следовательно, к оцениванию η^2 $(=\rho^2)$.

Будем рассматривать каждый столбец таблицы 26.7 как y-сечение. Пусть n_p обозначает число наблюдений в p-м сечении, $n = \sum n_p$; μ_p — среднее, а σ_p^2 — дисперсия y в этом сечении; μ — среднее, а σ_y^2 — дисперсия величины y. Предположим, что все измерения по y сделаны от значения k, являющегося

точкой дихотомии; это не ограничивает общности, так как ρ^2 и η^2 инвариантны относительно изменения начала отсчета. Тогда оценкой корреляционного отношения y по x (см. (26.40)) служит величина

$$\frac{\frac{1}{n}\sum_{p=1}^{q} n_p\mu_p^2 - \mu_y^2}{\sigma_y^2} = \frac{1}{n}\sum_{p=1}^{q}\frac{n_p\mu_p^2}{\sigma_p^2}\cdot\frac{\sigma_p^2}{\sigma_y^2} - \frac{\mu_y^2}{\sigma_y^2}. \qquad (26.68)$$

Но для двумерного нормального распределения $\eta^2 = \rho^2$ и (см. **16.23**)

$$\sigma_p^2/\sigma_y^2 = \mathsf{D}(y \mid x)/\sigma_y^2 = (1-\rho^2),$$

так что, заменяя в (26.68) σ_p^2/σ_y^2 на $(1-\rho^2)$, получаем уравнение

$$\rho^2 \approx \frac{1-\rho^2}{n}\sum_{p=1}^{q}\frac{n_p\mu_p^2}{\sigma_p^2} - \frac{\mu_y^2}{\sigma_y^2}. \qquad (26.69)$$

Решая его относительно ρ^2, находим оценку

$$r_\eta^2 = \frac{\frac{1}{n}\sum_{p=1}^{q} n_p\left(\frac{\mu_p}{\sigma_p}\right)^2 - \left(\frac{\mu_y}{\sigma_y}\right)^2}{1+\frac{1}{n}\sum_{p=1}^{q} n_p\left(\frac{\mu_p}{\sigma_p}\right)^2}. \qquad (26.70)$$

Эту оценку называют *бисериальным коэффициентом* η вследствие его сходства с корреляционным отношением. Придерживаясь нашего соглашения, по которому латинские буквы используются для статистик, выборочное значение этого коэффициента мы обозначаем r_η.

Использование выражения (26.70) оказывается возможным потому, что входящие в него величины могут быть оценены по выборочным данным. Из предположения о двумерной нормальности исходного распределения следует, что величина, согласно которой сделана дихотомия (в нашем примере — алкоголизм), может быть представлена некоторой вариантой, имеющей нормальное распределение, и что каждое y-сечение является дихотомией одномерного нормального распределения. Таким образом, отношения (μ_p/σ_p) и (μ_y/σ_y) могут быть оценены с помощью таблиц нормального интеграла. Например, в таблице 26.7 частоты алкоголиков и неалкоголиков, в колонке «Поджог» равны 50 и 43. Следовательно, относительная частота

алкоголиков равна $50/93 = 0{,}5376$, а нормальное отклонение, соответствующее этой частоте, есть, согласно таблице, 0,0944, что служит оценкой величины $|\mu_p/\sigma_p|$ для этого сечения.

Пример 26.12

Для данных таблицы 26.7 находим относительные частоты, оценки отношений $|\mu_p/\sigma_p|$ и $|\mu_y/\sigma_y|$ и значения n_p:

	Поджог	Изнасилование	Насильственные действия	Воровство	Изготовление фальшивых денег	Мошенничество	Всего
Алкоголики	0,5376	0,5867	0,5849	0,5582	0,5625	0,3043	0,5281
$\lvert\mu_p/\sigma_p\rvert$	0,0944	0,2190	0,2144	0,1463	0,1573	0,5119	$0{,}0704 = \lvert\mu_y/\sigma_y\rvert$
n_p	93	150	265	679	32	207	$1426 = n$

Теперь из (26.70) получаем

$$r_\eta^2 = \frac{\frac{1}{1426}\{93\,(0{,}0944)^2 + \ldots\} - (0{,}0704)^2}{1 + \frac{1}{1426}\{93\,(0{,}0944)^2 + \ldots\}} = 0{,}05456,$$

или

$$|r_\eta| = 0{,}234,$$

что можно принять, согласно нашим предположениям, за оценку коэффициента корреляции.

26.31 Как и в случае тетрахорического коэффициента r_t, выборочное распределение бисериального коэффициента r_η неизвестно. К. Пирсон (1917) нашел асимптотическое выражение для его выборочной дисперсии, но неизвестно, как велико должно быть n, чтобы этим выражением можно было пользоваться.

Нельзя ожидать, что r_t или r_η может быть очень эффективной оценкой для ρ, так как они используют слишком мало информации о переменных. Стоит помнить также, что предположение о двумерной нормальности исходного распределения (хотя оно не всегда делалось явно) было решающим для обоих методов. Без предположения о нормальности неизвестно в общем случае, что оценивают r_t и r_η.

26.32 Если в таблице $2 \times q$ классификация по q группам задана значениями некоторой варианты (в отличие от неупорядоченной классификации, как в таблице 26.7), то вместо η можно оценивать непосредственно ρ. Действительно, нам теперь

доступна дополнительная информация, которая позволяет оценить дисперсию σ_x^2 этой числовой варианты и ее средние μ_1, μ_2, относящиеся к каждой из частей дихотомии относительно y. Так как регрессия x по y линейна, то имеем (см. (26.12))

$$\mathbf{M}(x \mid y) - \mu_x = \rho \frac{\sigma_x}{\sigma_y}(y - \mu_y). \tag{26.71}$$

Как в **26.27**, можно найти такое k, чтобы

$$1 - F(k) = (2\pi)^{-1/2} \int_k^\infty \exp\left(-\frac{1}{2}u^2\right) du = \frac{n_1}{n_1 + n_2}, \tag{26.72}$$

где n_1 — общее количество индивидуумов, носящих один признак y-класса («более высокие» значения y), а n_2 — количество индивидуумов с другим признаком. Значение k является точкой дихотомии нормального распределения величины y.

Из (26.71) следует, что средние (y_i, μ_i) $(i = 1, 2)$ каждой части дихотомии будут лежать на линии регрессии (26.71). Таким образом, для части дихотомии, содержащей «более высокое» значение y (обозначим его y_1), имеем

$$\rho = \left(\frac{\mathbf{M}(x \mid y_1) - \mu_x}{\sigma_x}\right) \Big/ \left(\frac{y_1 - \mu_y}{\sigma_y}\right).$$

Следовательно, в качестве оценки для ρ можно взять отношение

$$\left(\frac{\bar{x}_1 - \bar{x}}{s_x}\right) \Big/ \left(\frac{y_1 - \mu_y}{\sigma_y}\right), \tag{26.73}$$

где $\bar{x}_1$, $\bar{x}$ — средние значения x при «более высоких» наблюдениях y и по всей таблице соответственно, а s_x^2 — выборочная дисперсия x по всей таблице. Знаменатель в (26.73), согласно (26.72), имеет вид

$$\frac{y_1 - \mu_y}{\sigma_y} = (2\pi)^{-1/2} \int_k^\infty u \exp\left(-\frac{1}{2}u^2\right) du \Big/ (2\pi)^{-1/2} \int_k^\infty \exp\left(-\frac{1}{2}u^2\right) du =$$

$$= (2\pi)^{-1/2} \exp\left(-\frac{1}{2}k^2\right) \Big/ \left(\frac{n_1}{n_1 + n_2}\right).$$

Если обозначить ординату плотности нормального распределения в точке k через z_k, то получим оценку для ρ

$$r_b = \left(\frac{\bar{x}_1 - \bar{x}}{s_x}\right) \frac{n_1}{n_1 + n_2} \frac{1}{z_k}. \tag{26.74}$$

Оценку, даваемую этим равенством, мы обозначили r_b, указывая индексом на ее название: r_b называется «бисериальным коэффициентом r».

Последнее равенство обычно записывается в более симметричной форме. Так как

$$\bar{x} = (n_1\bar{x}_1 + n_2\bar{x}_2)/(n_1 + n_2),$$

то разность $\bar{x}_1 - \bar{x}$ равна $n_2(\bar{x}_1 - \bar{x}_2)/(n_1 + n_2)$. Обозначая p отношение $n_1/(n_1 + n_2)$ и $q = 1 - p$, находим иной вид для (26.74):

$$r_b = \frac{\bar{x}_1 - \bar{x}_2}{s_x}\,\frac{pq}{z_k}. \qquad (26.75)$$

Пример 26.13 (из К. Пирсона, 1909)

Таблица 26.8 содержит распределение по возрасту кандидатов (всего 6156 человек), сдавших и не сдавших приемные экзамены в Лондонский университет в 1908, 1909 гг. Для двух старших возрастных групп указаны оценки среднего возраста.

Таблица 26.8

Возраст кандидата	Сдавшие	Не сдавшие	Всего
16	583	563	1146
17	666	980	1646
18	525	868	1393
19—21	383	814	1197
22—30 (среднее 25 лет)	214	439	653
Свыше 30 (среднее 33 года)	40	81	121
Всего	2411	3745	6156

Используя индекс 1 для успешно сдавших кандидатов, имеем

$$\bar{x}_1 = 18{,}4280.$$

По всем вместе кандидатам

$$\bar{x} = 18{,}7685, \qquad s_x^2 = (3{,}2850)^2.$$

Значение p равно $2411/6156 = 0{,}3917$.

Соотношение (26.72) дает $1 - F(k) = 0{,}3917$, и мы находим $k = 0{,}275$ и $z_k = 0{,}384$. Следовательно, из (26.74)

$$r_b = -\frac{0{,}3405}{3{,}2850}\cdot\frac{0{,}3917}{0{,}384} = -0{,}11.$$

Оцененная корреляция между возрастом и успехом невелика.

26.33 Как в случае r_t и r_η, предположение об исходной нормальности было решающим в выводе r_b. Распределение бисериального коэффициента r_b неизвестно, но Сопер (1914) получил выражение для его стандартной ошибки в случае нормальных

выборок

$$\mathrm{D}r_b \sim \frac{1}{n}\left[\rho^4 + \rho^2\left\{\frac{pqk^2}{z_k^2} + (2p-1)\frac{k}{z_k} - \frac{5}{2}\right\} + \frac{pq}{z_k^2}\right] \quad (26.76)$$

и показал, что хорошей аппроксимацией к (26.76) служит

$$\mathrm{D}r_b \sim \frac{1}{n}\left[r_b^2 - \frac{(pq)^{1/2}}{z_k}\right].$$

Позже подробное исследование коэффициента r_b было проведено Марицем (1953) и Тейтом (1955), которые показали, что в нормальных выборках он распределен асимптотически нормально со средним ρ и дисперсией (26.76). Кроме того, они рассмотрели оценку максимального правдоподобия для ρ при бисериальных данных. Оказалось, как можно было ожидать, что при фиксированном ρ дисперсия величины r_b минимальна, когда дихотомия производится в середине интервала дихотомизируемой переменной ($y = 0$). Если $\rho = 0$, то r_b является эффективной оценкой коэффициента ρ, но если $\rho^2 \to 1$, то эффективность r_b стремится к нулю. Тейт также табулировал формулу Сопера (26.76) для $\mathrm{D}r_b$. См. упражнения 26.10—26.12.

Точечно-бисериальная корреляция

26.34 Рассмотрим еще один коэффициент, *точечно-бисериальную корреляцию*, который обозначим через ρ_{pb}, а его выборочное значение — через r_{pb}. Предположим, что дихотомия относительно y не является разбиением нормального распределения, а определяется случайной величиной, принимающей только два значения. Поскольку речь идет о корреляции, то примем эти значения равными 1 и 0. Например, в таблице 26.8 предположение, что успех на экзамене есть дихотомия нормального распределения способностей сдавать экзамены, не является неправдоподобным. Но если бы y-дихотомия была относительно, скажем, пола, то это уже не было бы разумным предположением, и к подобной задаче необходим иной подход.

В действительности такая ситуация коренным образом отличается от рассматривавшейся до сих пор, так как теперь речь идет не об оценке коэффициента ρ двумерной нормальной совокупности: вместо этого мы имеем смешанный момент величины y, принимающей значения 0—1, и величины x. Если P есть истинная доля значений y с $y = 1$, а $Q = 1 - P$, то из теории биномиального распределения находим

$$\mathbf{M}(y) = P, \qquad \sigma_y^2 = PQ.$$

Следовательно, по определению

$$\rho_{pb} = \frac{\mu_{11}}{\sigma_x \sigma_y} = \frac{\mathbf{M}(xy) - P\mathbf{M}(x)}{\sigma_x (PQ)^{1/2}}.$$

Оценивая $\mathbf{M}(xy)$ через $m_{11} = \frac{1}{n_1 + n_2} \sum_{i=1}^{n_1} x_i$, $\mathbf{M}(x)$ через $\bar{x}$, σ_x через s_x и P через $p = \frac{n_1}{n_1 + n_2}$, получаем

$$r_{pb} = \frac{p\bar{x}_1 - p(p\bar{x}_1 + q\bar{x}_2)}{s_x (pq)^{1/2}} = \frac{(\bar{x}_1 - \bar{x}_2)(pq)^{1/2}}{s_x}. \qquad (26.77)$$

26.35 Коэффициент r_{pb} из (26.77) можно сравнить с бисериальным коэффициентом r_b, определенным в (26.75). Имеем

$$\frac{r_{pb}}{r_b} = \frac{z_k}{(pq)^{1/2}}. \qquad (26.78)$$

Рассматривая отношение Миллса (см. **5.22**), Тейт (1953) показал, что величина в правой части (26.78) не превосходит $(2/\pi)^{1/2}$ и, следовательно, значения коэффициентов будут, в общем случае, заметно различными.

Тейт (1954) показал, что r_{pb} распределен асимптотически нормально со средним ρ_{pb} и дисперсией

$$\mathbf{D}r_{pb} \sim \frac{(1 - \rho_{pb}^2)^2}{n}\left(1 - \frac{3}{2}\rho_{pb}^2 + \frac{\rho_{pb}^2}{4pq}\right), \qquad (26.79)$$

которая минимальна при $p = q = 1/2$.

Отвлекаясь от измерения корреляции, мы видим из (26.77), что в действительности в точечно-бисериальной ситуации мы сравниваем средние значения величины x, вычисленные по двум выборкам, а y-классификация состоит просто в различении двух выборок. На самом деле

$$\frac{r_{pb}^2}{1 - r_{pb}^2} = \frac{(\bar{x}_1 - \bar{x}_2)\frac{n_1 n_2}{n_1 + n_2}}{\sum (x_{1i} - \bar{x}_1)^2 + \sum (x_{2i} - \bar{x}_2)^2} = \frac{t^2}{n_1 + n_2 - 2}, \qquad (26.80)$$

где t — обычная t-статистика Стьюдента, используемая для сравнения средних двух нормальных совокупностей с равными дисперсиями (см. пример 23.8). Таким образом, если распределение величины x нормально для $y = 0, 1$, то точечно-бисериальный коэффициент является простым преобразованием t^2-статистики, которая может быть использована для проверки гипотез о нем.

26.36 Изложенное выше не исчерпывает всех возможных оценок коэффициента ρ двумерной нормальной совокупности по данным, классифицированным в таблицу с двумя входами. В главе 33 мы рассмотрим некоторые оценки, основанные на ранговых статистиках.

УПРАЖНЕНИЯ

26.1 Показать, что коэффициент корреляции для данных таблицы 26.2 равен 0,072. Проверить, что линии регрессии на рис. 26.2 определяются уравнениями

$$CC'\colon\ y = 0{,}0938x + 30{,}56; \qquad RR'\colon\ x = 0{,}0547y + 61{,}06.$$

26.2 Представляя двумерную функцию плотности в виде

$$f(x, y) = f(x)\, g(y \mid x),$$

так что j-й момент y-сечения при заданном x относительно начала координат равен

$$\mu_j'(y \mid x) = \int_{-\infty}^{\infty} y^j g(y \mid x)\, dy,$$

показать, что

$$\left[\frac{\partial^j \varphi(t, u)}{\partial u^j}\right]_{u=0} = i^j \int_{-\infty}^{\infty} e^{itx} f(x)\, \mu_j'(y \mid x)\, dx$$

(где φ — характеристическая функция распределения), откуда

$$f(x)\, \mu_j'(y \mid x) = \frac{(-i)^j}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \left[\frac{\partial^j \varphi}{\partial u^j}\right]_{u=0} dt.$$

Используя это, проверить, что двумерное нормальное распределение имеет линейные регрессии и постоянную условную дисперсию (является гомоскедастичным).

(Уикселл, 1934.)

26.3 Двумерное нормальное распределение дихотомизировано в некотором значении величины y. Известно, что дисперсия величины x для всего распределения равна σ_x^2, а для одной из частей дихотомии равна σ_1^2. Корреляция между x и y для этой части дихотомии равна c. Показать, что корреляция между x и y для всего распределения может быть оценена величиной r, где

$$r^2 = 1 - \frac{\sigma_1^2}{\sigma_x^2}(1 - c^2).$$

26.4 Показать, что если в предыдущем упражнении σ_y^2 — дисперсия величины y для всего распределения, а σ_2^2 — ее дисперсия в выбранной там части дихотомии, то коэффициент ρ для всего распределения может быть оценен с помощью

$$r^2 = \frac{c^2 \sigma_y^2}{\sigma_2^2 + c^2(\sigma_y^2 - \sigma_2^2)}.$$

26.5 Показать, что в отличие от тетрахорического r_t, бисериального r_η и точечно-бисериального r_{pb} коэффициентов, которые никогда не превосходят единицы, бисериальный коэффициент r_b может ее превосходить.

26.6 Доказать, что тетрахорический ряд (26.67) всегда сходится для $|\rho| < 1$.

26.7 Случайные величины $x_1, x_2, \ldots, x_n$ распределены так, что коэффициент корреляции между x_i и x_j равен ρ_{ij}. Все они имеют одинаковые дисперсии. Показать, что среднее значение ρ_{ij}, определяемое как

$$\bar{\rho} = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij}, \qquad i \neq j,$$

не может быть меньше, чем $-1/(n-1)$.

26.8 В условиях предыдущего упражнения показать, что $|\rho_{ij}|$, детерминант матрицы коэффициентов корреляции, неотрицателен. Вывести отсюда, что

$$\rho_{12}^2 + \rho_{13}^2 + \rho_{23}^2 \leqslant 1 + 2\rho_{12}\rho_{13}\rho_{23}.$$

26.9 Из (16.86) вывести, что для выборки из двумерной нормальной совокупности выборочное распределение величины b_2, коэффициента регрессии y по x, имеет дисперсию

$$\mathrm{D}b_2 = \frac{1}{n-3} \frac{\sigma_2^2}{\sigma_1^2} (1 - \rho^2),$$

а его асимметрия и эксцесс равны

$$\gamma_1 = 0, \qquad \gamma_2 = \frac{6}{n-5}.$$

26.10 Пусть $\psi\{(x-\mu)/\sigma, y\}$ обозначает двумерную нормальную плотность со средними значениями x и y, равными μ и 0, дисперсиями σ^2 и 1 соответственно и корреляцией ρ. Определим

$$\xi(x, \omega) = \int_{\omega}^{\infty} \psi \, dy, \qquad \eta(x, \omega) = \int_{-\infty}^{\omega} \psi \, dy.$$

Пусть z_i — случайная величина, принимающая значения 0 и 1, когда $y < \omega$ и $y \geqslant \omega$ соответственно. Показать, что для бисериальной таблицы функция правдоподобия может быть представлена в виде

$$L(x, y \mid \omega, \rho, \mu, \sigma) = \prod_{i=1}^{n} \left\{ z_i \xi\left(\frac{x_i - \mu}{\sigma}, \omega \right) + (1 - z_i)\, \eta\left(\frac{x_i - \mu}{\sigma}, \omega \right) \right\}.$$

Пусть ∂^2 обозначает частный дифференциал второго порядка относительно любой пары параметров. Показать, что

$$\mathbf{M}(\partial^2 \log L) = n\{1 - p(x)\}\, \mathbf{M}_0(\partial^2 \log \eta) + np(x)\, \mathbf{M}_1(\partial^2 \log \xi),$$

где

$$p(x) = \int_{x}^{\infty} (2\pi)^{-1/2} \exp\left(-\frac{1}{2} t^2 \right) dt,$$

а $\mathbf{M}_0$, $\mathbf{M}_1$ — условные математические ожидания по x при условии, что $y < \omega$, $y \geqslant \omega$ соответственно. Вывести отсюда, что матрица, обратная матрице дис-

персий для оценок максимального правдоподобия четырех параметров (порядок строк и столбцов такой же, как порядок параметров в ФП), имеет вид

$$\boldsymbol{V}^{-1}=\frac{n}{(1-\rho^2)}\begin{Vmatrix} a_0 & \dfrac{\rho\omega a_0-a_1}{1-\rho^2} & \dfrac{\rho a_0}{\sigma} & \dfrac{\rho a_1}{\sigma} \\ & \dfrac{a_2-2\rho\omega a_1+\rho^2\omega^2 a_0}{(1-\rho^2)^2} & \dfrac{\rho^2\omega a_0-\rho a_1}{\sigma(1-\rho^2)} & \dfrac{\rho^2\omega a_1-\rho a_2}{\sigma(1-\rho^2)} \\ & & \dfrac{1-\rho^2+\rho^2 a_0}{\sigma^2} & \dfrac{\rho^2 a_1}{\sigma^2} \\ & & & \dfrac{2(1-\rho^2)+\rho^2 a_2}{\sigma^2} \end{Vmatrix},$$

где

$$a_s=\int_{-\infty}^{\infty} x^s g(x,\omega,\rho)\,dx,$$

$$g(x,\omega,\rho)=(2\pi)^{-1/2}\exp\left(-\frac{1}{2}x^2\right)\varphi\left(\frac{\omega-\rho x}{(1-\rho^2)^{1/2}}\right)\varphi\left(\frac{\rho x-\omega}{(1-\rho^2)^{1/2}}\right)$$

и

$$\varphi(x)=(2\pi)^{-1/2}\exp\left(-\frac{1}{2}x^2\right)\Big/\{1-p(x)\}.$$

Обратив эту матрицу, получить асимптотическую дисперсию оценки максимального правдоподобия $\hat{\rho}_b$ в форме

$$\mathrm{D}\hat{\rho}_b=\frac{(1-\rho^2)^3}{n}\left\{\frac{\int_{-\infty}^{\infty} g\,dx}{\int_{-\infty}^{\infty} g\,dx\int_{-\infty}^{\infty} x^2 g\,dx-\left(\int_{-\infty}^{\infty} xg\,dx\right)^2}\right\}+\frac{\rho^2(1-\rho^2)}{n}.$$

(Тейт, 1955.)

26.11 В условиях упражнения 26.10 показать, что, когда $\rho=0$,

$$\mathrm{D}\hat{\rho}_b=\frac{2\pi p(\omega)\{1-p(\omega)\}}{n\exp(-k^2)}.$$

Сравнивая этот результат с асимптотической формулой (26.76), показать, что r_b при $\rho=0$ есть эффективная оценка.

(Тейт, 1955.)

26.12 В условиях упражнения 26.10 показать, что $n\,\mathrm{D}\hat{\rho}_b$ стремится к нулю, когда $|\rho|$ стремится к единице. Используя (26.76), доказать, что $n\,\mathrm{D}r_b$ не стремится к нулю и, следовательно, что r_b имеет нулевую эффективность вблизи $|\rho|=1$.

(Тейт, 1955; результаты упражнений 26.10—26.12 обобщены Хеннаном и Тейтом (1965) на многомерное нормальное распределение.)

26.13 Вывести равенства (26.19) и (26.20).

26.14 Обозначая l выборочный внутриклассовый коэффициент, а λ его теоретическое значение, показать, что точное распределение величины l

представляется в виде

$$dF \propto \frac{(1-l)^{p(k-1)/2-1}\{1+(k-1)\,l\}^{(p-3)/2}\,dl}{\{1-\lambda+\lambda(k-1)(1-l)\}^{(kp-1)/2}},$$

где l вычисляется по p семействам, каждое из которых содержит k членов, и что в случае $k=2$ оно приводится к виду

$$dF = \frac{\Gamma(p-1/2)}{\Gamma(p-1)(2\pi)^{1/2}}\operatorname{sech}^{p-1/2}(z-\xi)\exp\left\{-\frac{1}{2}(z-\xi)\right\}dz,$$

где $l=\operatorname{th} z$, $\lambda=\operatorname{th}\xi$. Вывести отсюда, что при $k=2$ величина $z-\xi$ имеет распределение, близкое к нормальному, со средним нуль и дисперсией $\frac{1}{n-3/2}$.

(Фишер, 1921c.)

26.15 Показать, что для проверки гипотезы $\rho=\rho_0$ в двумерной нормальной совокупности статистика отношения правдоподобия дается соотношением

$$l^{1/n} = \frac{(1-r^2)^{1/2}(1-\rho_0^2)^{1/2}}{(1-r\rho_0)}.$$

Следовательно, $l^{1/n}=(1-r^2)^{1/2}$, когда $\rho_0=0$, а когда $\rho_0\neq 0$, то

$$(1-\rho_0^2)^{-1/2}\,l^{1/n} = 1+r\rho_0+r^2\left(\rho_0^2-\frac{1}{2}\right)+\cdots$$

26.16 Доказать, что использование поправок Шеппарда для моментов всегда ведет к увеличению коэффициента корреляции.

26.17 Показать, что если x и y подвержены соответственно ошибкам наблюдений u, v, где u и v некоррелированы с x и y и между собой, то коэффициент корреляции делится («ослабляется») на

$$\left\{\left(1+\frac{\sigma_u^2}{\sigma_x^2}\right)\left(1+\frac{\sigma_v^2}{\sigma_y^2}\right)\right\}^{1/2}.$$

26.18 Пусть x_1, x_2, x_3 взаимно некоррелированы, имеют положительные средние и малые коэффициенты вариации V_i $(i=1, 2, 3)$.

Показать, что корреляция между x_1/x_3 и x_2/x_3 приближенно равна

$$\rho = \frac{V_3^2}{\{(V_1^2+V_3^2)(V_2^2+V_3^2)\}^{1/2}} > 0.$$

(Эту величину иногда называют «ложной» («spurious») корреляцией. Причина такого названия в некоррелированности исходных случайных величин, но это не совсем удачный термин.)

26.19 Пусть две двумерные нормальные совокупности имеют $\rho_1=\rho_2=\rho$, а остальные параметры не заданы. Показать, что оценка максимального правдоподобия для ρ есть

$$\hat{\rho} = \frac{n(1+r_1r_2) - \{n^2(1-r_1r_2)^2 - 4n_1n_2(r_1-r_2)^2\}^{1/2}}{2(n_1r_2+n_2r_1)},$$

где n_i, r_i — объемы выборок и коэффициенты корреляции $(i=1, 2)$, а $n=n_1+n_2$. Пусть $n_1=n_2=\frac{1}{2}n$. Показать, что если z_1, z_2 определены

соотношением (26.25) и

$$\hat{\zeta} = \frac{1}{2} \log\left(\frac{1+\hat{\rho}}{1-\hat{\rho}}\right),$$

то в точности

$$\hat{\xi} = \frac{1}{2}(z_1 + z_2).$$

26.20 Используя результат последнего упражнения, доказать, что критерий отношения правдоподобия для гипотезы $\rho_1 = \rho_2$, когда $n_1 = n_2$, основан на статистике

$$l^{1/n} = \operatorname{sech}\left\{\frac{1}{2}(z_1 - z_2)\right\}.$$

Следовательно, это есть взаимно однозначная функция от $z_1 - z_2$, статистики, предложенной в **26.19.**

(Брэнднер, 1933.)

26.21 В условиях упражнения 26.19 показать, что если $n_1 \neq n_2$, то МП-оценка величины ζ *приближенно* равна

$$\hat{\zeta} = \frac{1}{n}(n_1 z_1 + n_2 z_2)$$

и, следовательно, критерий ОП гипотезы $\rho_1 = \rho_2$ использует приближенную статистику

$$l = \left[\operatorname{sech}\left\{\frac{n_1}{n}(z_1 - z_2)\right\}\right]^{n_2}\left[\operatorname{sech}\left\{\frac{n_2}{n}(z_1 - z_2)\right\}\right]^{n_1},$$

снова являющуюся взаимно однозначной функцией от $z_1 - z_2$.

(Брэнднер, 1933.)

26.22 В задаче оценивания общего значения ρ в двух двумерных нормальных совокупностях показать, что оценка для ζ

$$\zeta^* = \frac{(n_1 - 3)z_1 + (n_2 - 3)z_2}{n_1 + n_2 - 6}$$

есть линейная комбинация z_1 и z_2 с минимальной дисперсией. Доказать, что при $n_1 \neq n_2$ она не совпадает с оценкой максимального правдоподобия для ρ, приведенной в упражнении 26.19.

26.23 Показать, что коэффициент корреляции ρ_{xy} между x и y удовлетворяет соотношениям

$$\rho_{xy}^2 = \frac{\mathbf{D}(\alpha_2 + \beta_2 x)}{\sigma_2^2} \equiv 1 - \frac{\mathbf{M}\{[y - (\alpha_2 + \beta_2 x)]^2\}}{\sigma_2^2},$$

и получить отсюда (26.18).

26.24 Полагая $z = \mathbf{M}(x|y)$, показать, что

$$\rho_{xz}^2 = \eta_1^2$$

и что

$$\rho_{yz}^2 = \rho_{xy}^2/\eta_1^2.$$

Вывести отсюда, что (26.18) влечет (26.46), и установить условия, при которых имеют место различные равенства в (26.46).

(М. Фреше опубликовал эти результаты в 1933—1935 гг.; см. Крускал (1958).)

ГЛАВА 27

ЧАСТНАЯ И МНОЖЕСТВЕННАЯ КОРРЕЛЯЦИЯ

27.1 В случае двух нормальных или почти нормальных величин коэффициент корреляции между ними может быть использован, как мы видели в **26.10**, в качестве меры взаимозависимости. Однако на практике при интерпретации «взаимозависимости» часто встречаются те же трудности, которые уже обсуждались в **26.4**, а именно: если одна величина коррелирована с другой, то это может быть всего лишь отражением того факта, что они обе коррелированы с некоторой третьей величиной или с совокупностью величин. Указанная возможность приводит нас к рассмотрению условных корреляций между двумя величинами при фиксированных значениях остальных величин. Это так называемые *частные корреляции*.

Если корреляция между двумя величинами уменьшается, когда мы фиксируем некоторую другую случайную величину, то это означает, что их взаимозависимость возникает частично через воздействие этой величины; если же частная корреляция равна нулю или очень мала, то мы делаем вывод, что их взаимозависимость целиком обусловлена этим воздействием. Наоборот, когда частная корреляция больше первоначальной корреляции между двумя величинами, то мы заключаем, что другие величины ослабляли связь, или, можно сказать, «маскировали» корреляцию. Но нужно помнить, что даже в последнем случае мы не имеем права предполагать наличие причинной связи: согласно рассуждениям из **26.4**, некоторая совершенно отличная от рассматриваемых в нашем анализе величина может быть источником этой корреляции. Как при обычной корреляции, так и при частных корреляциях предположение о причинности должно всегда иметь внестатистические основания.

27.2 В этой области статистики временами трудно достигнуть недвусмысленных и гибких обозначений без того, чтобы они были крайне громоздкими. Основываясь на системе обозначений Юла (1907), мы будем придерживаться среднего курса, но иногда от читателя потребуется терпение к индексам.

Как и в главе 26, мы попутно будем рассматривать также линейную регрессию, но основное рассмотрение задач регрессии отложим до главы 28.

Частная корреляция

27.3 Рассмотрим три величины, имеющие трехмерное нормальное распределение. Исключим вырожденный случай (см. **15.2**) и без потери общности, поскольку мы касаемся лишь корреляций, будем считать величины нормированными. Тогда их матрица рассеяния совпадает с матрицей их корреляций, которую назовем *корреляционной матрицей* и обозначим $\boldsymbol{C}$. Таким образом, если корреляция между x_i и x_j есть ρ_{ij}, то согласно (15.19) функция плотности имеет вид

$$f(x_1, x_2, x_3) = (2\pi)^{-3/2} |C|^{-1/2} \exp\left\{ -\frac{1}{2|C|} \sum_{i,j=1}^{3} C_{ij} x_i x_j \right\}, \qquad (27.1)$$

где C_{ij} — алгебраическое дополнение (i, j)-го элемента в симметричном корреляционном определителе

$$|C| = \begin{vmatrix} 1 & \rho_{12} & \rho_{13} \\ & 1 & \rho_{23} \\ & & 1 \end{vmatrix}. \qquad (27.2)$$

$C_{ij}/|C| = C^{ij}$ есть элемент матрицы, обратной к $\boldsymbol{C}$. Мы будем иногда записывать определитель или матрицу корреляций в таком виде, когда оставлено свободным место ниже главной диагонали, которое должно заполняться по симметрии.

Из (15.20) находим х. ф. этого распределения

$$\varphi(t_1, t_2, t_3) = \exp\left\{ -\frac{1}{2} \sum_{i,j=1}^{3} \rho_{ij} t_i t_j \right\}. \qquad (27.3)$$

27.4 Рассмотрим корреляцию между x_1 и x_2 при фиксированном значении x_3. Условное распределение x_1 и x_2 при заданном x_3 равно

$$g(x_1, x_2 \mid x_3) \propto$$

$$\propto \exp\left\{ -\frac{1}{2}(C^{11}x_1^2 + 2C^{12}x_1x_2 + C^{22}x_2^2 + 2C^{13}x_1x_3 + 2C^{23}x_2x_3) \right\} \propto$$

$$\propto \exp\left\{ -\frac{1}{2}[C^{11}(x_1 - \xi_1)^2 + 2C^{12}(x_1 - \xi_1)(x_2 - \xi_2) + C^{22}(x_2 - \xi_2)^2] \right\}, \qquad (27.4)$$

где

$$C^{11}\xi_1 + C^{12}\xi_2 = -C^{13}x_3, \qquad C^{12}\xi_1 + C^{22}\xi_2 = -C^{23}x_3.$$

Из (27.4) видно, что при заданном x_3 величины x_1 и x_2 имеют двумерное нормальное распределение с коэффициентом корреляции

$$\rho_{12.3} = -\frac{C^{12}}{(C^{11}C^{22})^{1/2}}.$$

Ясно, что $\rho_{12.3}$ не зависит от фиксируемого значения величины x_3. Кроме того, сокращая на общий множитель $|C|$, из (27.2) находим

$$\rho_{12.3} = -\frac{C_{12}}{(C_{11}C_{22})^{1/2}} = \frac{\rho_{12}-\rho_{13}\rho_{23}}{\{(1-\rho_{13}^2)(1-\rho_{23}^2)\}^{1/2}}. \tag{27.5}$$

$\rho_{12.3}$ называется частным коэффициентом корреляции между x_1 и x_2 при фиксированном x_3. Он симметричен относительно *первичных индексов* 1, 2. Его *вторичный индекс* 3 относится к переменной, которая фиксирована.

Хотя (27.5) выведено в предположении нормальности, мы теперь для любого исходного распределения *определим* частный коэффициент корреляции с помощью (27.5).

27.5 Аналогично, если мы имеем p-мерное невырожденное нормальное распределение и фиксируем $p-2$ случайных величины, то получаем частную корреляцию оставшихся двух (скажем, x_1 и x_2):

$$\rho_{12.34\ldots p} = \frac{-C_{12}}{(C_{11}C_{22})^{1/2}}, \tag{27.6}$$

где C_{ij} — алгебраическое дополнение для ρ_{ij} в определителе

$$|C| = \begin{vmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1p} \\ & 1 & \rho_{23} & \cdots & \rho_{2p} \\ & & 1 & \cdots & \rho_{3p} \\ & & & \cdots & \cdots \\ & & & & 1 \end{vmatrix}. \tag{27.7}$$

Подобно (27.5), (27.6) следует рассматривать как общее определение частного коэффициента корреляции между x_1 и x_2 при фиксированных $x_3, \ldots, x_p$.

27.6 Полезно рассмотреть ту же задачу с другой точки зрения. Обозначим $f(x_1, \ldots, x_k | x_{k+1}, \ldots, x_p)$ условную совместную плотность распределения величин $x_1, \ldots, x_k$, когда $x_{k+1}, \ldots, x_p$ фиксированы, а $g(x_{k+1}, \ldots, x_p)$ — совместное маргинальное распределение $x_{k+1}, \ldots, x_p$.

Совместная х. ф. всех p величин есть

$$\varphi(t_1, \ldots, t_p) =$$

$$= \int \cdots \int f(x_1, \ldots, x_k | x_{k+1}, \ldots, x_p)\, g(x_{k+1}, \ldots, x_p) \times$$

$$\times \exp\left(\sum_{j=1}^{p} it_j x_j\right) dx_1 \ldots dx_p =$$

$$= \int \cdots \int \varphi_k(t_1, \ldots, t_k | x_{k+1}, \ldots, x_p)\, g(x_{k+1}, \ldots, x_p) \times$$

$$\times \exp\left(\sum_{j=k+1}^{p} it_j x_j\right) dx_{k+1} \ldots dx_p,$$

где $\varphi_k(t_1, \ldots, t_k | x_{k+1}, \ldots, x_p)$ — условная совместная х. ф. для $x_1, \ldots, x_k$. Из многомерной теоремы обращения (4.17) следует, что

$$\varphi_k g = \frac{1}{(2\pi)^{p-k}} \int \ldots \int \varphi(t_1, \ldots, t_p) \exp\left(- \sum_{j=k+1}^{p} it_j x_j \right) dt_{k+1} \ldots dt_p. \tag{27.8}$$

Если в (27.8) положить $t_1 = t_2 = \ldots = t_k = 0$, то из равенства φ_k единице получаем

$$g = \frac{1}{(2\pi)^{p-k}} \int \ldots \int \varphi(0, \ldots, 0, t_{k+1}, \ldots, t_p) \times$$

$$\times \exp\left(- \sum_{j=k+1}^{p} it_j x_j \right) dt_{k+1} \ldots dt_p. \tag{27.9}$$

Следовательно, после деления (27.8) на (27.9) находим

$$\varphi_k = \frac{\int \ldots \int \varphi(t_1, \ldots, t_p) \exp\left(- \sum_{j=k+1}^{p} it_j x_j \right) dt_{k+1} \ldots dt_p}{\int \ldots \int \varphi(0, \ldots, 0, t_{k+1}, \ldots, t_p) \exp\left(- \sum_{j=k+1}^{p} it_j x_j \right) dt_{k+1} \ldots dt_p}. \tag{27.10}$$

Этот общий результат вытекает из теоремы Бартлетта (1938).

Предположим теперь, что наши p величин имеют многомерное нормальное распределение. Тогда, используя их х. ф. (15.20), преобразуем подинтегральную функцию числителя в (27.10):

$$\exp\left(- \frac{1}{2} \sum_{l,\, j=1}^{p} \rho_{lj} t_l t_j - \sum_{j=k+1}^{p} it_j x_j \right) =$$

$$= \exp\left(- \frac{1}{2} \sum_{l,\, j=1}^{k} \rho_{lj} t_l t_j \right) \exp\left(- \frac{1}{2} \sum_{l,\, j=k+1}^{p} \rho_{lj} t_l t_j \right) \times$$

$$\times \exp\left(- \sum_{l=1}^{k} \sum_{j=k+1}^{p} \rho_{lj} t_l t_j \right) \exp\left(- \sum_{j=k+1}^{p} it_j x_j \right) =$$

$$= \exp\left(- \frac{1}{2} \sum_{l,\, j=1}^{k} \rho_{lj} t_l t_j \right) \exp\left(- \frac{1}{2} \sum_{l,\, j=k+1}^{p} \rho_{lj} t_l t_j \right) \times$$

$$\times \exp\left\{ - \sum_{j=k+1}^{p} it_j \left(x_j - i \sum_{l=1}^{k} \rho_{lj} t_l \right) \right\}. \tag{27.11}$$

Теперь интеграл относительно $t_{k+1}, \ldots, t_p$ от двух последних множителей в правой части (27.11) является обратным преобразованием многомерной нормальной х. ф. величин $x_{k+1}, \ldots, x_p$, причем x_j отсчитывается от значения $i \sum_{m=1}^{k} \rho_{jm} t_m$. Это изменение начал координат не влияет на корреляции. Если обозначить $\boldsymbol{D}$ корреляционную матрицу величин $x_{k+1}, \ldots, x_p$, то с точностью до постоянного множителя интеграл от (27.11) будет равен

$$\exp\left(-\frac{1}{2}\sum_{l,\ j=1}^{k}\rho_{lj}t_l t_j\right)\times$$

$$\times\exp\left\{-\frac{1}{2}\sum_{l,\ j=k+1}^{p}D^{lj}\left(x_l - i\sum_{m=1}^{k}\rho_{lm}t_m\right)\left(x_j - i\sum_{m=1}^{k}\rho_{jm}t_m\right)\right\}.$$

Учитывая сказанное, из (27.10) имеем

$$\varphi_k(t_1, \ldots, t_k \mid x_{k+1}, \ldots, x_p) =$$

$$= \exp\left(-\frac{1}{2}\sum_{l,\ j=1}^{k}\rho_{lj}t_l t_j\right)\exp\left\{-\frac{1}{2}\sum_{l,\ j=k+1}^{p}D^{lj}\left(x_l - i\sum_{m=1}^{k}\rho_{lm}t_m\right)\times\right.$$

$$\left.\times\left(x_j - i\sum_{m=1}^{k}\rho_{jm}t_m\right) + \frac{1}{2}\sum_{l,\ j=k+1}^{p}D^{lj}x_l x_j\right\}. \qquad (27.12)$$

Таким образом, если σ'_{uv} обозначает ковариацию между x_u и x_v в условном распределении величин $x_1, \ldots, x_k$, а σ_{uv} — их безусловную ковариацию, то, сравнивая в (27.12) коэффициенты при $t_u t_v$, находим

$$\sigma'_{uv} = \sigma_{uv} - \sum_{l,\ j=k+1}^{p} D^{lj}\rho_{lu}\rho_{jv}. \qquad (27.13)$$

Это выражение получено в предположении, что исходные величины нормированы. Если теперь отказаться от нормировки, так что x_i будет иметь дисперсию σ_i^2, то каждое ρ заменится на соответствующее ему σ, D^{lj} — на $D^{lj}/(\sigma_l\sigma_j)$, и мы получим более общую форму соотношения (27.13):

$$\sigma'_{uv} = \sigma_{uv} - \sum_{l,\ j=k+1}^{p} D^{lj}\sigma_{lu}\sigma_{jv}/(\sigma_l\sigma_j). \qquad (27.14)$$

Равенство (27.14) не зависит от фиксируемых значений $x_{k+1}, \ldots, x_p$.

Если обозначить безусловную $(k \times k)$-матрицу рассеяния $\{\sigma_{uv}\}$ через $\boldsymbol{A}$, $(k \times (p-k))$-матрицу $\{\sigma_{lu}\}$ через $\boldsymbol{B}'$ и $((p-k) \times (p-k))$-матрицу рассеяния, из которой $\boldsymbol{D}$ получается в результате нормировки, через $\boldsymbol{E}$, то (27.14) утверждает, что условная матрица рассеяния равна

$$\{\sigma'_{uv}\} = \boldsymbol{A} - \boldsymbol{B}'\boldsymbol{E}^{-1}\boldsymbol{B}.$$

27.7 В частности, если зафиксировать только одну переменную, скажем x_p, то $D^{pp} = 1$, и условная ковариация (27.14) тогда равна

$$\sigma'_{uv} = \sigma_{uv} - \sigma_{pu}\sigma_{pv}/\sigma_p^2 = \sigma_u\sigma_v(\rho_{uv} - \rho_{up}\rho_{vp}). \qquad (27.15)$$

При $u = v$ из (27.15) находим условную дисперсию u:

$$\sigma'^2_u = \sigma_u^2(1 - \rho_{up}^2).$$

Из двух последних формул получаем условный коэффициент корреляции

$$\rho_{uv.p} = \frac{\rho_{uv} - \rho_{up}\rho_{vp}}{\{(1-\rho_{up}^2)(1-\rho_{vp}^2)\}^{1/2}}$$

того же вида, что и (27.5).

Если зафиксируем все переменные, кроме двух, скажем x_1 и x_2, то из (27.14) будем иметь

$$\rho_{12.34\ldots p} = \frac{\rho_{12} - \sum_{l,\,j=3}^{p} D^{lj}\rho_{l1}\rho_{j2}}{\left\{\left(1 - \sum_{l,\,j=3}^{p} D^{lj}\rho_{l1}\rho_{j1}\right)\left(1 - \sum_{l,\,j=3}^{p} D^{lj}\rho_{l2}\rho_{j2}\right)\right\}^{1/2}}. \qquad (27.16)$$

Рассматривая (27.7), находим, что минор элемента ρ_{12}, а именно

$$\begin{vmatrix} \rho_{21} & \rho_{23} & \rho_{24} & \cdots & \rho_{2p} \\ \rho_{31} & 1 & \rho_{34} & \cdots & \rho_{3p} \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \rho_{p1} & \rho_{p3} & \rho_{p4} & \cdots & 1 \end{vmatrix} = \begin{vmatrix} \rho_{21} & \rho_{32} & \cdots & \rho_{p2} \\ \rho_{31} & & & \\ \cdot & & & \\ \cdot & & \boldsymbol{D} & \\ \cdot & & & \\ \rho_{p1} & & & \end{vmatrix},$$

может быть разложен по его первой строке и столбцу в виде

$$\rho_{21}|D| - \sum_{l,\,j=3}^{p} D^{lj}\rho_{l1}\rho_{j2},$$

и аналогично для миноров элементов ρ_{11}, ρ_{22}. Таким образом, (27.16) представимо в форме

$$\rho^2_{12.34\ldots p} = \frac{C_{12}^2}{C_{11}C_{22}},$$

что вновь совпадает с (27.6).

Частные линейные регрессии

27.8 Обобщим теперь соотношения линейной регрессии из **26.7** на случай p величин. Для p совместно нормальных величин x_i с нулевыми средними и дисперсиями σ_i^2 математическое ожидание величины x_1 при условии, что $x_2, \ldots, x_p$ фиксированы, как видно из выражения в экспоненте распределения, равно

$$\frac{\mathbf{M}(x_1 \mid x_2, \ldots, x_p)}{\sigma_1} = -\sum_{j=2}^{p} \frac{C_{1j}}{C_{11}} \frac{x_j}{\sigma_j}. \qquad (27.17)$$

Коэффициент регрессии x_1 по x_j при фиксированных остальных $p-2$ величинах будем обозначать $\beta_{1j.23\ldots j-1,\, j+1\ldots p}$ или, короче, $\beta_{1j.q_j}$, где q символизирует совокупность величин, отличных от указанных первичными индексами, а индекс у q служит для различения этих совокупностей. Коэффициенты $\beta_{1j\cdot q}$ называются *частными коэффициентами регрессии.*

Следовательно, мы имеем

$$\mathbf{M}(x_1 \mid x_2, \ldots, x_p) = \beta_{12.q_2}x_2 + \beta_{13.q_3}x_3 + \ldots + \beta_{1p.q_p}x_p. \qquad (27.18)$$

Сравнивая (27.18) с (27.17), получаем для многомерного нормального случая

$$\beta_{1j.q_j} = -\frac{\sigma_1}{\sigma_j}\frac{C_{1j}}{C_{11}}. \qquad (27.19)$$

Аналогично, коэффициент регрессии x_j по x_1 при фиксированных остальных переменных есть

$$\beta_{j1.q_j} = -\frac{\sigma_j}{\sigma_1}\frac{C_{j1}}{C_{jj}}. \qquad (27.20)$$

Таким образом, поскольку $C_{1j} = C_{j1}$, то из (27.6), (27.19) и (27.20) получаем

$$\rho^2_{1j.q_j} = \frac{C_{1j}^2}{C_{11}C_{jj}} = \beta_{1j.q_j}\beta_{j1.q_j} \qquad (27.21)$$

— очевидное обобщение соотношения (26.17). Соотношения (27.19) и (27.20) показывают, что коэффициент $\beta_{1j.q_j}$ не сим-

метричен относительно x_1 и x_j, как и следовало ожидать от коэффициента зависимости. Подобно (27.5) и (27.6), (27.19) и (27.20) являются *определениями* частных коэффициентов регрессии в общем случае.

Ошибки относительно линейной регрессии

27.9 Назовем *ошибкой* *) порядка $(p-1)$ величину

$$x_{1.2\ldots p} = x_1 - \mathbf{M}(x_1 \mid x_2, \ldots, x_p).$$

Ее среднее равно нулю, а дисперсия равна

$$\sigma^2_{1.2\ldots p} = \mathbf{M}(x^2_{1.2\ldots p}) = \mathbf{M}\left[\{x_1 - \mathbf{M}(x_1 \mid x_2, \ldots, x_p)\}^2\right],$$

так что $\sigma^2_{1.2\ldots p}$ является *дисперсией ошибки* величины x_1 относительно регрессии. Из (27.18) немедленно получаем

$$\sigma^2_{1.2\ldots p} = \mathbf{M}\left[\left\{x_1 - \sum_{j=2}^{p} \beta_{1j.q_j} x_j\right\}^2\right] = \tag{27.22}$$

$$= \mathbf{M}\left[x_1\left(x_1 - \sum_{j=2}^{p} \beta_{1j.q_j} x_j\right) - \sum_{j=2}^{p} \beta_{1j.q_j} x_j \left(x_1 - \sum_{j=2}^{p} \beta_{1j.q_j} x_j\right)\right]. \tag{27.23}$$

Если брать математические ожидания в два этапа, фиксируя вначале $x_2, \ldots, x_p$, то условное математическое ожидание от второго члена в (27.23) будет равно, согласно (27.18), нулю. Таким образом,

$$\sigma^2_{1.2\ldots p} = \mathbf{M}\left\{x_1^2 - \sum_{j=2}^{p} \beta_{1j.q_j} x_1 x_j\right\} = \sigma_1^2 - \sum_{j=2}^{p} \beta_{1j.q_j} \sigma_{1j}. \tag{27.24}$$

Дисперсия ошибки (27.24) не зависит от фиксируемых значений $x_3, \ldots, x_p$, если только от них не зависят коэффициенты $\beta_{1j.q_j}$. В этом случае условное распределение величины x_1 называется *гомоскедастическим* (*homoscedastic*) (или *гетероскедастическим* (*heteroscedastic*) в противном случае). Это постоянство дисперсии ошибки делает интерпретацию регрессий и корреляций более простой. Например, в нормальном случае условные дисперсии и ковариации, полученные при фиксировании множества величин, не зависят от значений, в которых последние фиксированы (см. (27.14)). В других случаях при интерпретациях мы должны надлежащим образом учитывать обнаруженную гетероскедастичность, тогда, возможно, частные коэффициенты регрессии лучше всего рассматривать как показатели зависимости, *усредненные* по всевозможным значениям фиксированных величин.

*) В литературе эту величину часто называют «остатком» (residual), но мы будем проводить различие между *ошибками* (errors) относительно линейных регрессий в генеральной совокупности и *остатками*, возникающими при подгонке регрессий к выборочным данным.

Соотношения между дисперсиями, регрессиями и корреляциями различных порядков

27.10 Если даны p величин, то мы можем изучать корреляцию между любыми двумя из них, когда среди оставшихся зафиксированы значения произвольного подмножества величин. Аналогично, можно интересоваться регрессией произвольной величины относительно любого подмножества из оставшихся величин. С возрастанием p число всевозможных коэффициентов становится очень большим. Если некоторый коэффициент содержит k вторичных индексов, то говорят, что он имеет порядок k. Так, порядок $\rho_{12.34}$ равен 2, порядок $\rho_{12.3}$ — единице, порядок ρ_{12} — нулю, тогда как $\beta_{12.678}$ имеет порядок 3, а $\sigma^2_{1.2678}$ — порядок 4. В наших нынешних обозначениях коэффициенты линейной регрессии из последней главы β_1 и β_2 должны быть записаны в виде β_{12} и β_{21} соответственно. Они имеют порядок нуль, как и обычная дисперсия σ^2.

В **27.4** и **27.7** мы уже видели, как любой коэффициент корреляции первого порядка может быть выражен через коэффициенты нулевого порядка. Теперь будут получены более общие результаты такого сорта для коэффициентов всех типов.

27.11 Из (27.24) и (27.19) имеем

$$\sigma^2_{1.2\ldots p} = \sigma_1^2 + \sum_{j=2}^{p} \frac{\sigma_1}{\sigma_j} \frac{C_{1j}}{C_{11}} \sigma_{1j}, \tag{27 25}$$

откуда

$$\frac{\sigma^2_{1.2\ldots p}}{\sigma_1^2} = 1 + \frac{1}{C_{11}} \sum_{j=2}^{n} C_{1j}\rho_{1j} = 1 + \frac{1}{C_{11}}(|C| - C_{11}) = \frac{|C|}{C_{11}}.$$

Пользуясь символом q, введенным в **27.8**, получаем

$$\sigma^2_{1.q} = \sigma_1^2 \frac{|C|}{C_{11}}, \tag{27.26}$$

и аналогично, если 1 заменить любым другим индексом.

Точно таким же путем можно получить более общий результат

$$\operatorname{cov}\left(x_{l.q_l}, x_{m.q_m}\right) = \sigma_l \sigma_m \frac{|C|}{C_{lm}}, \tag{27.27}$$

который сводится к (27.26) при $l = m$. Соотношение (27.27) применимо в случае, когда вторичные индексы одной величины включают в себя первичные индексы другой. Если, с другой стороны, оба множества вторичных индексов не содержат l и m, то

обозначим через r общее множество вторичных индексов. Ковариация двух ошибок $x_{l.r}$, $x_{m.r}$ связана с их корреляцией и дисперсиями соотношениями, естественно обобщающими определения (26.10), (26.11) и (26.17), а именно:

$$\left.\begin{aligned} \operatorname{cov}(x_{l.r}, x_{m.r})/\sigma^2_{r.m} &= \beta_{lm.r}, \\ \operatorname{cov}(x_{l.r}, x_{m.r})/\sigma^2_{l.r} &= \beta_{ml.r}, \\ \operatorname{cov}(x_{l.r}, x_{m.r})/(\sigma_{l.r}\sigma_{m.r}) &= \rho_{lm.r}, \end{aligned}\right\} \tag{27.28}$$

что согласуется с уже найденным соотношением (27.21). Присоединяя множество индексов r к обеим величинам x_l, x_m, мы попросту должны сделать то же самое со всеми их коэффициентами.

27.12 Теперь можно использовать (27.26) для получения соотношения между дисперсиями ошибок различных порядков. Обозначая $|D|$ корреляционный определитель всех величин, кроме x_2, имеем из (27.26)

$$\sigma^2_{1.q-2} = \sigma^2_1 \frac{|D|}{D_{11}}$$

(где индекс $q-2$ обозначает множество q без x_2) и

$$\sigma^2_{1.q} = \sigma^2_1 \frac{|C|}{C_{11}},$$

откуда

$$\frac{\sigma^2_{1.q}}{\sigma^2_{1.q-2}} = \frac{D_{11}}{C_{11}} \cdot \frac{|C|}{|D|}. \tag{27.29}$$

По определению $|D| = C_{22}$, а согласно обобщенной теореме Якоби об определителях

$$\begin{vmatrix} C_{11} & C_{12} \\ C_{12} & C_{22} \end{vmatrix} = |C| D_{11}, \tag{27.30}$$

так как D_{11} является дополнительным минором для

$$\begin{vmatrix} \rho_{11} & \rho_{12} \\ \rho_{12} & \rho_{22} \end{vmatrix}$$

в $\boldsymbol{C}$. Таким образом, используя (27.30), получаем из (27.29)

$$\frac{\sigma^2_{1.q}}{\sigma^2_{1.q-2}} = \frac{\begin{vmatrix} C_{11} & C_{12} \\ C_{12} & C_{22} \end{vmatrix}}{C_{11}C_{22}} = 1 - \frac{C^2_{12}}{C_{11}C_{22}} \tag{27.31}$$

или, учитывая (27.6), находим

$$\sigma^2_{1.q} = \sigma^2_{1.q-2}(1 - \rho^2_{12.q}). \tag{27.32}$$

Соотношение (27.32) является обобщением двумерного результата, приведенного в упражнении 26.23, который в наших нынешних обозначениях может быть представлен в виде

$$\sigma_{2.1}^2 = \sigma_2^2(1 - \rho_{12}^2).$$

Этот результат мы также встречали в (16.46) в связи с двумерным нормальным распределением.

27.13 Соотношение (27.32) дает нам возможность выразить дисперсию ошибки порядка $(p-1)$ через дисперсию ошибки и коэффициент корреляции порядка $(p-2)$. Если мы теперь вновь воспользуемся (27.32) для того, чтобы выразить $\sigma_{1.q-2}^2$, то тем же путем найдем, что

$$\sigma_{1.q-2}^2 = \sigma_{1.q-2-3}^2(1 - \rho_{13.q-2}^2).$$

Применяя последовательно (27.32) и записывая более полно индексы, получаем

$$\sigma_{1.2\ldots p}^2 = \sigma_1^2(1-\rho_{1p}^2)(1-\rho_{1(p-1).p}^2)(1-\rho_{1\,(p-2).(p-1)\,p}^2)\cdots \\ \cdots(1-\rho_{12.3\ldots p}^2). \qquad (27.33)$$

В (27.33), очевидно, не играет роли порядок вторичных индексов у $\sigma_{1.23\ldots p}$; мы их можем переставить так, как пожелаем. Например, для простоты можно написать

$$\frac{\sigma_{1.2\ldots p}^2}{\sigma_1^2} = (1-\rho_{12}^2)(1-\rho_{13.2}^2)(1-\rho_{14.23}^2)\cdots \\ \cdots(1-\rho_{1p.23\ldots(p-1)}^2) = \frac{|C|}{C_{11}} \qquad (27.34)$$

в силу (27.26). В (27.34) индексы, отличные от 1, допускают перестановку. Соотношение (27.34) позволяет нам выразить дисперсию ошибки порядка s через дисперсию ошибки нулевого порядка и s коэффициентов корреляции, порядок которых принимает значения от нуля до $(s-1)$.

27.14 Перейдем теперь к коэффициентам регрессии. Перепишем (27.15) для ковариации между x_1 и x_2 при фиксированном x_p:

$$\sigma_{12.p} = \sigma_{12} - \sigma_{1p}\sigma_{p2}/\sigma_p^2.$$

Присоединяя повсюду индексы $3,\ldots,(p-1)$, имеем

$$\sigma_{12.3\ldots p} = \sigma_{12.3\ldots(p-1)} - \frac{\sigma_{1p.3\ldots(p-1)}\,\sigma_{p2.3\ldots(p-1)}}{\sigma_{p.3\ldots(p-1)}^2}. \qquad (27.35)$$

Используя определение (27.28) коэффициента регрессии как отношения ковариации к дисперсии, т. е.

$$\beta_{ij.k} = \sigma_{ij.k}/\sigma_{j.k}^2,$$

и обозначая через r множество $3, \ldots, (p-1)$, находим из (27.35)

$$\beta_{12.pr}\sigma^2_{2.pr} = \beta_{12.r}\sigma^2_{2.r} - \beta_{1p.r}\beta_{p2.r}\sigma^2_{2.r},$$

или

$$\beta_{12.pr} = \frac{\sigma^2_{2.r}}{\sigma^2_{2.pr}}(\beta_{12.r} - \beta_{1p.r}\beta_{p2.r}). \tag{27.36}$$

Если в (27.36) положить $x_1 \equiv x_2$, то получим

$$\sigma^2_{2.pr} = \sigma^2_{2.r}(1 - \beta_{2p.r}\beta_{p2.r}) = \sigma^2_{2.r}(1 - \rho^2_{2p.r}), \tag{27.37}$$

другую форму соотношения (27.32). Таким образом, из (27.36) и (27.37) имеем

$$\beta_{12.pr} = \frac{\beta_{12.r} - \beta_{1p.r}\beta_{p2.r}}{1 - \beta_{2p.r}\beta_{p2.r}}. \tag{27.38}$$

Это и есть требуемая формула для выражения коэффициента регрессии через некоторые коэффициенты следующего более низкого порядка. Повторно применяя (27.38), найдем представление любого коэффициента регрессии в терминах коэффициентов нулевого порядка.

Наконец, используя (27.21), из (27.38) получаем соотношение

$$\rho_{12.pr} = (\beta_{12.pr}\beta_{21.pr})^{1/2} = \frac{\rho_{12.r} - \rho_{1p.r}\rho_{2p.r}}{\{(1 - \rho^2_{1p.r})(1 - \rho^2_{2p.r})\}^{1/2}}, \tag{27.39}$$

обобщающее (27.5) путем присоединения множества индексов r.

Приближенные частные линейные регрессии

27.15 В нашем изложении, начиная с **27.8**, мы занимались точно линейными регрессионными зависимостями типа (27.18). Так же как в **26.8**, рассмотрим теперь вопрос подгонки регрессионных соотношений этого типа к наблюденным совокупностям, регрессии которых почти никогда не бывают точно линейными. С помощью тех же рассуждений мы приходим к методу наименьших квадратов. Мы выбираем поэтому $\beta_{1j.q_j}$ так, чтобы минимизировать сумму квадратов уклонений n наблюдений от подгоняемой регрессии:

$$\sum_{i=1}^{n}\left(x_{1i} - \sum_{j=2}^{p}\beta_{1j.q_j}x_{ji}\right)^2, \tag{27.40}$$

где иксы измеряются от своих средних значений и предполагается $n > p$. Согласно (19.12) решение имеет вид

$$\boldsymbol{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{x}_1, \tag{27.41}$$

где матрица $\boldsymbol{X}$ составлена из наблюдений над $p-1$ величинами $x_2, \ldots, x_p$, а $\boldsymbol{x}_1$ — вектор наблюдений величины x_1. Соотношение (27.41) можно переписать в виде

$$\boldsymbol{\beta} = (n\boldsymbol{V}_{p-1})^{-1}(n\boldsymbol{M}) = \boldsymbol{V}_{p-1}^{-1}\boldsymbol{M}, \tag{27.42}$$

где $\boldsymbol{V}_{p-1}$ — матрица рассеяния для $x_2, \ldots, x_p$, а $\boldsymbol{M}$ — вектор ковариаций между x_1 и x_j $(j = 2, \ldots, p)$. Таким образом,

$$\beta_{1j \cdot q_j} = \frac{1}{|V_{p-1}|} \sum_{l=2}^{p} (V_{p-1})_{jl}\, \sigma_{1l}. \tag{27.43}$$

Поскольку $|V_{p-1}|$ есть минор V_{11} матрицы рассеяния $\boldsymbol{V}$ всех p величин, то $(V_{p-1})_{jl}$ представляет собой дополнительный минор для

$$\begin{vmatrix} \sigma_1^2 & \sigma_{1l} \\ \sigma_{1j} & \sigma_{jl} \end{vmatrix}$$

в $\boldsymbol{V}$, так что сумма в правой части (27.43) является алгебраическим дополнением для $(-\sigma_{1j})$ в $\boldsymbol{V}$. Поэтому (27.43) представляется в виде

$$\beta_{1j \cdot q_j} = -\frac{V_{1j}}{V_{11}} = -\frac{\sigma_1}{\sigma_j}\frac{C_{1j}}{C_{11}}. \tag{27.44}$$

Соотношение (27.44) совпадает с (27.19). Таким образом, как и в **26.8**, мы приходим к заключению, что аппроксимация по методу наименьших квадратов дает те же коэффициенты регрессии, что и в случае точной линейной регрессии.

Из этого следует, что все результаты данной главы остаются в силе, когда для наблюденных совокупностей мы подгоняем регрессии по методу наименьших квадратов.

Выборочные коэффициенты

27.16 Если при заданной выборке объема n мы подгоняем регрессии по методу наименьших квадратов, то все полученные ранее соотношения будут выполняться и для выборочных коэффициентов. Следуя нашему обычному соглашению, будем использовать r вместо ρ, b вместо β и s^2 вместо σ^2 для того, чтобы отличать выборочные коэффициенты от их генеральных значений. Коэффициенты b определяются с помощью минимизации величины, аналогичной (27.40),

$$ns_{1.23\ldots p}^2 = \sum_{i=1}^{n} \left(x_{1i} - \sum_{j=2}^{p} b_{1j \cdot q_j} x_{ji} \right)^2. \tag{27.45}$$

Как и в (27.21), имеем

$$r^2_{ij.k} = b_{ij.k} b_{ji.k};$$

точно так же выполняются соотношения, аналогичные (27.34), (27.38) и (27.39).

Для нахождения b_{1j} приравняем к нулю производные от (27.45) по b_{1j}. Получим $p - 1$ уравнений

$$\sum_{i=1}^{n} x_{ji} \left(x_{1i} - \sum_{j=2}^{p} b_{1j.q_j} x_{ji} \right) = 0, \qquad j = 2, 3, \ldots, p,$$

которые можно записать в виде

$$\sum x_j x_{1.23\ldots p} = 0, \qquad j = 2, 3, \ldots, p, \tag{27.46}$$

где суммирование происходит по всем n наблюдениям. Величина $x_{1.23\ldots p}$ есть *остаток* при подгонке регрессии (см. **27.9**). Из (27.46) следует, что

$$\sum x^2_{1.q} = \sum x_{1.q} (x_1 - \sum b_{1j.q_j} x_j) = \sum x_{1.q} x_1, \tag{27.47}$$

и аналогично

$$\sum x_{1.r} x_{2.r} = \sum x_{1.r} x_2 = \sum x_1 x_{2.r}, \tag{27.48}$$

где r есть произвольное множество общих вторичных индексов. Соотношения, подобные (27.47) и (27.48), выполняются как для выборочных остатков, так и для ошибок генеральной совокупности, но они найдут применение главным образом в выборочных задачах, поэтому они и выражены в терминах остатков. Упражнение 27.5 дает наиболее общее правило для того, чтобы опускать общие вторичные индексы при суммировании произведений остатков.

Оценивание генеральных коэффициентов

27.17 Так же как в случае регрессий и корреляций нулевого порядка в предыдущей главе, мы можем использовать выборочные коэффициенты в качестве оценок их истинных значений. Если рассматриваемая регрессия линейна, то из теории метода наименьших квадратов, изложенной в главе 19, известно, что любой коэффициент b является несмещенной оценкой соответствующего β, а $\frac{n}{n-(p-1)} s^2_{1.23\ldots p}$ — несмещенной оценкой для $\sigma^2_{1.23\ldots p}$. Однако r не является несмещенной оценкой для ρ: в **26.15—17** мы видели, что даже для коэффициента нулевого порядка в нормальном случае r не является несмещенной оценкой

коэффициента ρ, но что модификация (26.34) или (26.35) является несмещенной оценкой. Результат, который будет получен в 27.22, позволит нам оценить любой частный коэффициент корреляции аналогично нормальному случаю.

Геометрическая интерпретация частной корреляции

27.18 Из наших результатов вытекает, что вся совокупность частных регрессий, корреляций, а также дисперсий или ковариаций ошибок или остатков полностью определяется дисперсиями и корреляциями или же дисперсиями и регрессиями нулевого порядка. Интересно взглянуть на этот результат с геометрической точки зрения. Предположим, что у нас есть n наблюдений над p ($<n$) случайными величинами

$$x_{11}, \ldots, x_{1p}; \quad x_{21}, \ldots, x_{2p}; \ldots; \quad x_{n1}, \ldots, x_{np}.$$

Рассмотрим *n-мерное* (евклидово) выборочное пространство. Наблюдениям $x_{1k}, \ldots, x_{nk}$ над k-й случайной величиной в этом пространстве будет соответствовать одна точка. Следовательно, имеется p точек, по одной на каждую величину. Обозначим эти точки $Q_1, Q_2, \ldots, Q_p$. Предположим, что иксы отсчитываются от своих средних, и пусть точка P является началом координат.

Величину $n\sigma_l^2$ можно тогда интерпретировать как квадрат длины вектора, соединяющего точку Q_l (с координатами $x_{1l}, \ldots, x_{nl}$) с P. Аналогично, ρ_{lm} можно представлять себе как косинус угла Q_lPQ_m, ибо

$$\rho_{lm} = \frac{\sum_{j=1}^{n} x_{jl}x_{jm}}{\left(\sum_{j=1}^{n} x_{jl}^2 \sum_{j=1}^{n} x_{jm}^2\right)^{1/2}},$$

а это есть формула косинуса угла между PQ_l и PQ_m.

Наш результат тогда состоит в том, что все соотношения, связывающие p точек в n-мерном пространстве, могут быть выражены в терминах длин векторов PQ_i и углов между ними. Таким образом, теория частной корреляции и регрессии формально тождественна с тригонометрией некоторой совокупности точек в n-мерном пространстве.

27.19 Читатель, предпочитающий геометрический взгляд на рассматриваемые вопросы, без труда может переформулировать предыдущие уравнения, используя тригонометрическую терминологию. Мы укажем только наиболее важные результаты, требующиеся для дальнейших выборочных исследований.

Заметим прежде всего, что p точек Q_i и точка P определяют (исключая, возможно, вырожденный случай) подпространство

размерности p в n-мерном пространстве. Рассмотрим точку $Q_{1.2\ldots p}$, координатами которой являются n остатков $x_{1.2\ldots p}$. Согласно (27.46) вектор $PQ_{1.2\ldots p}$ ортогонален к каждому из векторов $PQ_2, \ldots, PQ_p$ и, следовательно, к подпространству размерности $(p-1)$, натянутому на $P, Q_2, \ldots, Q_p$.

Рассмотрим теперь векторы остатков $Q_{1.r}$, $Q_{2.r}$ и $Q_{p.r}$, где r заменяет вторичные индексы 3, 4, ..., $(p-1)$. Косинус угла между $Q_{1.r}$ и $Q_{2.r}$, скажем θ, равен $\rho_{12.r}$, и каждый из них ортогонален к подпространству, натянутому на P, Q_3, ..., $Q_{(p-1)}$. Пусть на рис. 27.1 точка M будет основанием перпендикуляра, опущенного из $Q_{1.r}$ на $PQ_{p.r}$, а $Q'_{2.r}$ — такой точкой на $PQ_{2.r}$, что вектор $Q'_{2.r}M$ перпендикулярен к $PQ_{p.r}$. Тогда $MQ_{1.r}$ и $MQ'_{2.r}$ ортогональны к пространству, натянутому на P, $Q_3, \ldots$, Q_p, и косинус угла между ними, скажем φ, равен $\rho_{12.rp}$. Таким образом, для того чтобы $\rho_{12.rp}$ выразить через $\rho_{12.r}$, мы должны выразить φ через θ или угол между векторами $PQ_{1.r}$ и $PQ'_{2.r}$ — через угол между их проекциями на гиперплоскость, перпендулярную к $PQ_{p.r}$. Опустим теперь для удобства штрих у $Q'_{2.r}$. По теореме Пифагора

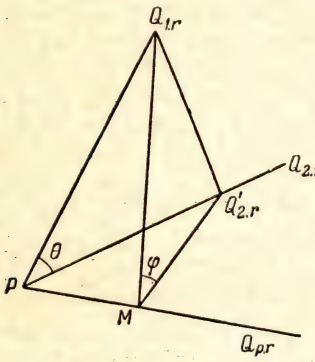


Рис. 27.1. Геометрия частной корреляции.

$$(Q_{1.r}Q_{2.r})^2 = PQ_{1.r}^2 + PQ_{2.r}^2 - 2PQ_{1.r}\cdot PQ_{2.r}\cos\theta =$$
$$= MQ_{1.r}^2 + MQ_{2.r}^2 - 2MQ_{1.r}\cdot MQ_{2.r}\cos\varphi.$$

Далее,

$$PQ_{1.r}^2 = PM^2 + MQ_{1.r}^2 \quad \text{и} \quad PQ_{2.r}^2 = PM^2 + MQ_{2.r}^2.$$

Следовательно, находим

$$MQ_{1.r}\cdot MQ_{2.r}\cos\varphi = PQ_{1.r}\cdot PQ_{2.r}\cos\theta - PM^2$$

или

$$\frac{MQ_{1.r}}{PQ_{1.r}}\cdot\frac{MQ_{2.r}}{PQ_{2.r}}\cos\varphi = \cos\theta - \frac{PM}{PQ_{1.r}}\cdot\frac{PM}{PQ_{2.r}}. \qquad (27.49)$$

Отношения $\dfrac{MQ_{1.r}}{PQ_{1.r}}$ и $\dfrac{PM}{PQ_{1.r}}$ являются синусом и косинусом угла между $PQ_{p.r}$ и $PQ_{1.r}$. Косинус угла между $PQ_{1.r}$ и $PQ_{p.r}$ равен $\rho_{1p.r}$, и, следовательно, $\dfrac{PM}{PQ_{1.r}} = \rho_{1p.r}$, $\dfrac{MQ_{1.r}}{PQ_{1.r}} = (1-\rho_{1p.r}^2)^{1/2}$. Та-

кой же результат получается при замене индекса 1 на 2. Итак, согласно (27.49)

$$\rho_{12.rp} = \frac{\rho_{12.r} - \rho_{1p.r}\rho_{2p.r}}{\{(1-\rho^2_{1p.r})(1-\rho^2_{2p.r})\}^{1/2}}, \qquad (27.50)$$

что вновь совпадает с (27.39). Таким образом, мы видим, что выражение частного коэффициента корреляции через коэффициент более низкого порядка можно представлять себе как проектирование некоторого угла в выборочном пространстве на подпространство, ортогональное к переменной, фиксированной только в исходном коэффициенте более высокого порядка.

Вычисление коэффициентов

27.20 Там, где имеются только три или четыре случайных величины, мы можем вычислять частные корреляции и регрессии, исходя непосредственно из коэффициентов нулевого порядка, пользуясь соответствующими формулами из числа полученных выше. Если же присутствует большое число переменных, то удобно систематизировать арифметику в форме определителей. Действительно, нам нужно вычислить все миноры корреляционной матрицы $\boldsymbol{C}$, и затем, подставляя их в формулы (27.6), (27.19) и (27.26), мы получим коэффициенты корреляции и регрессии, а также дисперсии остатков (или ошибок) всех порядков. Ныне электронные вычислительные машины становятся широко доступными и можно избежать утомительных ручных вычислений.

Для малых p полезны таблицы таких величин, как $1-\rho^2$, $(1-\rho^2)^{1/2}$ и $\{(1-\rho_1^2)(1-\rho_2^2)\}^{-1/2}$. Полезны также тригонометрические таблицы. Например, при заданном ρ можно найти $\theta = \arccos\rho$ и, следовательно, $\sin\theta = (1-\rho^2)^{1/2}$, $\operatorname{cosec}\theta = (1-\rho^2)^{-1/2}$ и так далее.

The Kelley Statistical Tables (Harvard U. P., 1948) содержит значения $(1-\rho^2)^{1/2}$ для $\rho = 0{,}0001\,(0{,}0001)\,0{,}9999$.

Два следующих примера представляют как интерпретационный, так и вычислительный интерес.

Пример 27.1

Исследуя для некоторой области Англии влияние погоды на урожаи, Хукер (1907) нашел следующие средние, стандартные отклонения и корреляции между урожаем сена (x_1) в центнерах *) на акр, весенним количеством осадков (x_2) в дюймах

*) См. примечание на стр. 388. (*Прим. ред.*)

и накопленной за весну температурой выше 42° F (x_3) по данным за 20 лет:

$$\mu_1 = 28{,}02, \quad \sigma_1 = 4{,}42, \quad \rho_{12} = +0{,}80,$$
$$\mu_2 = 4{,}91, \quad \sigma_2 = 1{,}10, \quad \rho_{13} = -0{,}40,$$
$$\mu_3 = 594, \quad \sigma_3 = 85, \quad \rho_{23} = -0{,}56.$$

Основной интерес здесь представляет вопрос о влиянии погоды на величину урожаев, и поэтому мы рассмотрим только регрессию x_1 относительно двух других случайных величин. Из корреляций нулевого порядка выясняется, что урожай и осадки имеют положительную корреляцию, а урожай и накопленная весенняя температура — отрицательную. Как же интерпретировать последний результат? Может быть, высокая температура неблагоприятно влияет на урожайность или, возможно, отрицательная корреляция является следствием того, что высокая температура влечет уменьшение количества дождей, так что польза тепла более чем устраняется вредным воздействием засухи?

Для того чтобы пролить некоторый свет на этот вопрос, вычислим частные корреляции. Из (27.5) находим

$$\rho_{12.3} = \frac{\rho_{12} - \rho_{13}\rho_{23}}{\{(1-\rho_{13}^2)(1-\rho_{23}^2)\}^{1/2}} = \frac{0{,}80 - (-0{,}40)(-0{,}56)}{\{(1-0{,}40^2)(1-0{,}56^2)\}^{1/2}} = 0{,}759.$$

Аналогично,

$$\rho_{13.2} = 0{,}097, \quad \rho_{23.1} = -0{,}436.$$

Далее нам потребуются регрессии и дисперсии ошибок. Имеем

$$\beta_{12.3} = \frac{\operatorname{cov}(x_{1.3}, x_{2.3})}{\mathrm{D}x_{2.3}} = \rho_{12.3}\frac{\sigma_{1.3}}{\sigma_{2.3}}.$$

Это соотношение, однако, содержит величины $\sigma_{1.3}$ и $\sigma_{2.3}$, которые самостоятельного интереса не представляют. Их вычисления можно избежать, если заметить, что из (27.33) следуют равенства

$$\sigma_{1.23} = \sigma_{1.3}(1-\rho_{12.3}^2)^{1/2}, \quad \sigma_{2.13} = \sigma_{2.3}(1-\rho_{12.3}^2)^{1/2}, \quad (27.51)$$

так что

$$\beta_{12.3} = \rho_{12.3}\frac{\sigma_{1.23}}{\sigma_{2.13}}. \quad (27.52)$$

Стандартные отклонения $\sigma_{1.23}$ и $\sigma_{2.13}$ представляют некоторый интерес и могут быть вычислены по (27.33). Получаем

$$\sigma_{1.23} = \sigma_1\{(1-\rho_{12}^2)(1-\rho_{13.2}^2)\}^{1/2} = \sigma_1\{(1-\rho_{13}^2)(1-\rho_{12.3}^2)\}^{1/2}$$

— два выражения, дающие возможность взаимной проверки. Из первого имеем

$$\sigma_{1.23} = 4{,}42\,\{(1 - 0{,}8^2)(1 - 0{,}097^2)\}^{1/2} = 2{,}64.$$

Аналогично,

$$\sigma_{2.13} = 0{,}594, \qquad \sigma_{3.12} = 70{,}1.$$

Таким образом,

$$\beta_{12.3} = 0{,}759 \cdot \frac{2{,}64}{0{,}594} = 3{,}37.$$

Совершенно так же находим

$$\beta_{13.2} = 0{,}00364.$$

Уравнение регрессии x_1 относительно x_2 и x_3 представляется тогда в виде

$$x_1 - 28{,}02 = 3{,}37\,(x_2 - 4{,}91) + 0{,}00364\,(x_3 - 594).$$

Это уравнение показывает, что при увеличении количества осадков урожай возрастает и что при увеличении температуры урожай также возрастает, когда прочие факторы постоянны. Оно дает нам возможность изолировать воздействие осадков от воздействия температуры и изучать каждое из них по отдельности. Положительность коэффициента $\beta_{13.2}$ означает, что существует положительная связь между урожаем и температурой, когда влияние осадков устранено. Частные корреляции говорят то же самое. Несмотря на отрицательность ρ_{13}, коэффициент $\rho_{13.2}$ положителен (хотя и мал) и указывает на то, что отрицательное значение ρ_{13} является следствием сложного воздействия фактора осадков.

В предыдущих вычислениях не использовались определители, но при желании их можно применить. Определитель (27.2) есть

$$|C| = \begin{vmatrix} 1 & 0{,}80 & -0{,}40 \\ 0{,}80 & 1 & -0{,}56 \\ -0{,}40 & -0{,}56 & 1 \end{vmatrix} = 0{,}2448,$$

$$|C_{11}| = \begin{vmatrix} 1 & -0{,}56 \\ -0{,}56 & 1 \end{vmatrix} = 0{,}6864,$$

что при подстановке, например, в (27.34) дает полученный ранее результат

$$\sigma_{1.23} = \sigma_1 \left(\frac{|C|}{C_{11}}\right)^{1/2} = 2{,}64.$$

Пример 27.2

В исследованиях преступности в 16 больших городах США Огберн (1935) нашел корреляцию между степенью преступности (x_1), измеряемой числом известных преступлений на тысячу жителей, и церковным членством (x_5), измеряемым числом членов церкви старше 13 лет на 100 человек населения старше 13 лет. Она оказалась равной —0,14. Очевидный вывод состоит в том, что религиозность удерживает от преступления. Рассмотрим это более внимательно.

Пусть x_2 — процент мужского населения, x_3 — процент мужчин иностранного происхождения среди всего населения, x_4 — число детей до пятилетнего возраста на 1000 замужних женщин в возрасте от 15 до 44 лет. Огберн вычислил значения

$$\rho_{12} = +0{,}44, \quad \rho_{23} = +0{,}25, \quad \rho_{34} = +0{,}44,$$
$$\rho_{13} = -0{,}34, \quad \rho_{24} = -0{,}19, \quad \rho_{35} = +0{,}33,$$
$$\rho_{14} = -0{,}31, \quad \rho_{25} = -0{,}35, \quad \rho_{45} = +0{,}85.$$
$$\rho_{15} = -0{,}14,$$

С помощью этих и других приведенных в его работе данных можно показать, что регрессия величины x_1 относительно остальных четырех переменных имеет вид

$$x_1 - 19{,}9 = 4{,}51\,(x_2 - 49{,}2) - 0{,}88\,(x_3 - 30{,}2) - \\ - 0{,}072\,(x_4 - 4814) + 0{,}63\,(x_5 - 41{,}6),$$

а некоторые из частных корреляций равны

$$\rho_{15.3} = -0{,}03, \quad \rho_{15.4} = +0{,}25, \quad \rho_{15.34} = +0{,}23.$$

Из уравнения регрессии замечаем теперь, что при фиксированных остальных факторах величины x_1 и x_5 положительно связаны, т. е. церковное членство находится в положительной связи с преступностью. Что же маскирует этот эффект, давая отрицательную корреляцию в коэффициенте нулевого порядка ρ_{15}?

Заметим, прежде всего, что если устранить влияние величины x_3, процента иностранцев, то корреляция между преступностью и церковным членством будет близка к нулю. Эта же корреляция положительна, когда устранено воздействие x_4, числа маленьких детей. Если зафиксировать сразу x_3 и x_4, то корреляция вновь положительна. Из уравнения регрессии в действительности выясняется, что высокий процент иностранцев и большая доля детей отрицательно связаны со степенью преступности. Оба эти фактора положительно коррелированы с церковным членством (иностранные иммигранты являются в основном католиками и более многодетны). Во всей совокупно-

сти эти корреляции покрывают положительную связь преступности с церковным членством. Кажущаяся отрицательная связь церковного членства с преступностью, видимо, обусловлена бóльшим уважением к закону среди иностранных иммигрантов и тем фактом, что они более ревностные церковники.

Более полное обсуждение задачи читатель может найти в работе Огберна.

Выборочные распределения частных коэффициентов корреляции и регрессии в нормальном случае

27.21 Рассмотрим теперь выборочные распределения частных коэффициентов корреляции и регрессии в нормальном случае.

Для больших выборок с очевидными изменениями могут быть использованы стандартные ошибки, соответствующие коэффициентам нулевого порядка (см. **26.13**). Обозначая m множество вторичных индексов, имеем из (26.24)

$$\mathrm{D}r_{12.m} = \frac{1}{n}(1 - \rho^2_{12.m})^2, \tag{27.53}$$

а из (26.30) и (27.32)

$$\mathrm{D}b_{12.m} = \frac{1}{n}\frac{\sigma^2_{1.m}}{\sigma^2_{2.m}}(1 - \rho^2_{12.m}) = \frac{1}{n}\frac{\sigma^2_{1.2m}}{\sigma^2_{2.m}}. \tag{27.54}$$

Доказательство соотношений (27.53) и (27.54) с помощью прямых методов главы 10 очень утомительно. Однако они непосредственно следуют из того факта, что совместное распределение любых двух ошибок $x_{1.m}$ и $x_{2.m}$ является двумерным нормальным с коэффициентом корреляции $\rho_{12.m}$. Таким образом, как отметил Юл (1907), выборочные корреляции и регрессии между соответствующими остатками имеют, по крайней мере для большой выборки, то же распределение, что и коэффициент нулевого порядка. В **27.22** мы увидим, что для выборки объема n точное распределение величины $r_{12.m}$ таково же, как у коэффициента корреляции нулевого порядка, основанного на $n - d$ наблюдениях, где d есть число вторичных индексов в m. Однако, поскольку (27.53) и (27.54) верны лишь до членов порядка n^{-1}, нам нет нужды учитывать эту малую поправку.

27.22 Рассмотрим теперь геометрическое представление из **27.18—19.** Предположим, что у нас есть три вектора PQ_1, PQ_2, PQ_3, представляющие собой n наблюдений над x_1, x_2, x_3. Как мы видели в **27.19**, коэффициент частной корреляции $r_{12.3}$ является косинусом угла между PQ_1 и PQ_2, спроектированного на подпространство, ортогональное к PQ_3, размерность которого $(n - 1)$. Если сделать ортогональное преобразование (т. е.

поворот системы координат), то корреляции, являясь косинусами углов, останутся неизменными. Более того, если n исходных наблюдений над тремя величинами были независимыми, то таковыми останутся и n наблюдений над ортогонально преобразованными величинами. (Этот факт представляет собой обобщение результата, приведенного в примерах 11.2, 11.3 и в **15.27** для независимых переменных x_1, x_2, x_3, а его доказательство предоставляется читателю в виде упражнения 27.7; геометрически это очевидно из радиальной симметрии нормированного многомерного нормального распределения.) Выбирая PQ_3 в качестве одной из новых координатных осей при ортогональном преобразовании, сразу видим, что $r_{12.3}$ имеет такое же распределение, как коэффициент нулевого порядка, основанный на $n-1$ независимых наблюдениях. Повторяя это рассуждение, получаем, что распределение коэффициента корреляции порядка d, основанного на n наблюдениях, то же, что у коэффициента нулевого порядка, вычисленного по $n-d$ наблюдениям: каждый вторичный индекс влечет проектирование в выборочном пространстве на ортогональное дополнение к соответствующей переменной и потерю одной степени свободы. Приведенный результат принадлежит Фишеру (1924a).

С этим уточнением результаты предыдущей главы могут быть немедленно применены к частным корреляциям. Если d мало по сравнению с n, то распределение частных корреляций, когда n возрастает, по существу такое же, как у коэффициентов нулевого порядка, что подтверждает приближение (27.53) для стандартной ошибки.

Для частных коэффициентов регрессии, кроме того, получаем, что распределение (16.86) коэффициента нулевого порядка продолжает иметь место с заменой n на $n-d$, когда повсюду присоединено множество вторичных индексов m. В частности, статистика Стьюдента (26.38) для регрессии x_1 по x_2 превращается в

$$t=(b_{12.m}-\beta_{12.m})\left\{\frac{s^2_{2.m}(n-d-2)}{s^2_{1.m}-b^2_{12.m}s^2_{2.m}}\right\}^{1/2} \tag{27.55}$$

с $n-d-2$ степенями свободы. Если множество m содержит все $p-2$ оставшихся переменных, то число степеней свободы равно $n-p$. Поскольку коэффициенты регрессии зависят как от углов, так и от расстояний (дисперсий) в выборочном пространстве, то распределение величины b_{12}, в отличие от случая коэффициента r, непосредственно не сохраняется при проектировании (с уменьшением лишь числа степеней свободы): статистики $s^2_{1.m}$, $s^2_{2.m}$ в (27.55) осуществляют поправки, учитывающие изменение расстояний при проектировании.

Множественный коэффициент корреляции

27.23 Дисперсия ошибки величины x_1 относительно ее регрессии (27.18) по остальным переменным есть $\sigma^2_{1.2\ldots p}$; она была определена в **27.9**. Введем теперь *множественный коэффициент корреляции* *) $\boldsymbol{R}_{1(2\ldots p)}$ между x_1 и $x_2, \ldots, x_p$, полагая

$$1 - \boldsymbol{R}^2_{1(2\ldots p)} = \sigma^2_{1.2\ldots p}/\sigma^2_1. \tag{27.56}$$

Согласно (27.56) и (27.34) $0 \leqslant \boldsymbol{R}^2 \leqslant 1$. Определим коэффициент $\boldsymbol{R}$ как корень квадратный из $\boldsymbol{R}^2$; он всегда неотрицателен. Коэффициент $\boldsymbol{R}$, очевидно, несимметричен относительно своих индексов и представляет собой на самом деле меру *зависимости* x_1 от $x_2, \ldots, x_p$.

Для оправдания его названия покажем, что он действительно является коэффициентом корреляции. Согласно **27.9** имеем

$$\sigma^2_{1.2\ldots p} = \mathsf{M}(x^2_{1.2\ldots p}), \tag{27.57}$$

а аналог (27.47) для генеральной совокупности дает

$$\mathsf{M}(x^2_{1.2\ldots p}) = \mathsf{M}(x_1 x_{1.2\ldots p}). \tag{27.58}$$

Так как $\mathsf{M}(x_{1.2\ldots,p}) = 0$, то из (27.57) и (27.58) следует, что

$$\sigma^2_{1.2\ldots p} = \mathsf{D}(x_{1.2\ldots p}) = \operatorname{cov}(x_1, x_{1.2\ldots p}). \tag{27.59}$$

Если теперь рассмотреть корреляцию между x_1 и ее условным математическим ожиданием

$$\mathsf{M}(x_1 \mid x_2, \ldots, x_p) = x_1 - x_{1.2\ldots p},$$

то мы найдем, что она равна

$$\frac{\operatorname{cov}(x_1, x_1 - x_{1.2\ldots p})}{\{\mathsf{D}x_1 \mathsf{D}(x_1 - x_{1.2\ldots p})\}^{1/2}} = \frac{\mathsf{D}x_1 - \operatorname{cov}(x_1, x_{1.2\ldots p})}{[\mathsf{D}x_1 \{\mathsf{D}x_1 + \mathsf{D}x_{1.2\ldots p} - 2\operatorname{cov}(x_1, x_{1.2\ldots p})\}]^{1/2}}.$$

Подставляя сюда (27.59), получаем согласно (27.56)

$$\frac{\sigma^2_1 - \sigma^2_{1.2\ldots p}}{\{\sigma^2_1(\sigma^2_1 - \sigma_{1.2\ldots p})\}^{1/2}} = \left\{\frac{\sigma^2_1 - \sigma^2_{1.2\ldots p}}{\sigma^2_1}\right\}^{1/2} = \boldsymbol{R}_{1(2\ldots p)}. \tag{27.60}$$

Таким образом, $\boldsymbol{R}_{1(2\ldots p)}$ является обычным коэффициентом корреляции между x_1 и условным математическим ожиданием

*) Мы используем жирное $\boldsymbol{R}$ для коэффициента генеральной совокупности, а позднее будем использовать обычное заглавное R для обозначения соответствующего выборочного коэффициента. В соответствии с нашим соглашением следовало бы использовать греческую заглавную букву «ро» для обозначения коэффициента генеральной совокупности, но она напоминает заглавное P, что может вызвать путаницу.

$\mathbf{M}(x_1|x_2, \ldots, x_p)$. Поскольку сумма квадратов отклонений (и, следовательно, их среднее значение $\sigma^2_{1.2\ldots p}$) минимизируется при нахождении регрессии по методу наименьших квадратов, которая тождественна $\mathbf{M}(x_1|x_2, \ldots, x_p)$ (см. **27.15**), то из (27.60) следует, что $R_{1(2\ldots p)}$ представляет собой корреляцию между x_1 и «наиболее подходящей» линейной комбинацией величин $x_2, \ldots, x_p$. Никакая другая линейная функция от $x_2, \ldots, x_p$ не будет более коррелирована с x_1.

27.24 Из (27.56) и (27.34) получаем соотношения

$$1 - R^2_{1(2\ldots p)} = \frac{|C|}{C_{11}} = (1-\rho^2_{12})(1-\rho^2_{13.2}) \ldots (1-\rho^2_{1p.23\ldots(p-1)}), \tag{27.61}$$

выражающие множественный коэффициент корреляции либо через корреляционный определитель, либо через частные корреляции. Поскольку в (27.34) допустима перестановка индексов, отличных от 1, то из (27.61) немедленно вытекает (так как каждый множитель в правой части лежит в интервале (0, 1)), что

$$1 - R^2_{1(2\ldots p)} \leqslant 1 - \rho^2_{1j.s},$$

где $\rho_{1j.s}$ — произвольный частный или нулевого порядка коэффициент, содержащий 1 среди первичных индексов. Таким образом,

$$R_{1(2\ldots p)} \geqslant |\rho_{1j.s}|, \tag{27.62}$$

т. е. множественный коэффициент корреляции не меньше, чем абсолютная величина любого коэффициента корреляции с таким же первичным индексом. Если $R_{1(2\ldots p)} = 0$, то из этого следует, что и все соответствующие $\rho_{1j.s} = 0$. В таком случае величина x_1 полностью некоррелирована со всеми остальными величинами. С другой стороны, если $R_{1(2\ldots p)} = 1$, то по крайней мере один из коэффициентов $\rho_{1j.s}$ должен быть равен 1 для того, чтобы правая часть в (27.61) обращалась в нуль. В этом случае соотношение (27.56) показывает, что $\sigma^2_{1.2\ldots p} = 0$, т. е. все точки в распределении величины x_1 лежат на линии регрессии и x_1 является строго линейной функцией от $x_2, \ldots, x_p$.

Таким образом, коэффициент $R_{1(2\ldots p)}$ есть мера линейной зависимости величины x_1 от $x_2, \ldots, x_p$.

27.25 До сих пор мы рассматривали множественный коэффициент корреляции между x_1 и всеми остальными случайными величинами. Но, очевидно, можно также изучать множественную корреляцию между x_1 и любым подмножеством. Итак, определим

$$R^2_{1(s)} = 1 - \frac{\sigma^2_{1.s}}{\sigma^2_1} \tag{27.63}$$

для произвольного множества индексов s. Из (27.34) непосредственно вытекает, что

$$\sigma^2_{1.s} \leqslant \sigma^2_{1.r}, \tag{27.64}$$

где r — любое подмножество из s. Следовательно, дисперсия ошибки не может возрастать при добавлении последующих случайных величин. Таким образом, из (27.63) и (27.64) получаем соотношения вида

$$R^2_{1\,(2)} \leqslant R^2_{1\,(23)} \leqslant R^2_{1\,(234)} \leqslant \ldots \leqslant R^2_{1\,(2\ldots p)}, \tag{27.65}$$

выражающие тот факт, что коэффициент множественной корреляции никогда нельзя уменьшить путем расширения множества величин, относительно которых измеряется зависимость x_1.

В частности, при $p = 2$ из (27.61) находим

$$R^2_{1\,(2)} = \rho^2_{12}, \tag{27.66}$$

так что $R_{1(2)}$ равен абсолютной величине обычного коэффициента корреляции между x_1 и x_2.

Геометрическая интерпретация множественной корреляции

27.26 Можно интерпретировать $R_{1(2\ldots p)}$ в геометрических терминах пунктов **27.18—19.** Рассмотрим вначале интерпретацию регрессии (27.18), вычисленной по методу наименьших квадратов. Согласно **27.23** она является линейной функцией от $x_2, \ldots, x_p$, минимизирующей сумму квадратов (27.40). Таким образом, мы выбираем вектор PV в $(p-1)$-мерном подпространстве, натянутом на $P, Q_2, \ldots, Q_p$, который минимизирует расстояние Q_1V, т. е. который минимизирует угол между PQ_1 и PV. В силу (27.60) $R_{1(2\ldots p)}$ является косинусом этого минимального угла. Но последнее значит, что $R_{1(2\ldots p)}$ есть косинус угла между PQ_1 и самим $(p-1)$-мерным подпространством, потому что в противном случае угол не был бы минимальным.

Если $R_{1(2\ldots p)} = 0$, то вектор PQ_1 ортогонален к $(p-1)$-мерному подпространству, так что величина x_1 некоррелирована с $x_2, \ldots, x_p$ и с любой линейной функцией от них. С другой стороны, если $R_{1(2\ldots p)} = 1$, то PQ_1 лежит в $(p-1)$-мерном подпространстве, так что x_1 является строго линейной функцией от $x_2, \ldots, x_p$. Эти результаты получены в **27.24**.

Данная геометрическая интерпретация будет полезна при нахождении распределения выборочного коэффициента $R_{1(2\ldots p)}$ в нормальном случае. Она представляет собой прямое обобщение представления, использованного в **16.24** для получения распределения обычного коэффициента корреляции r, который согласно (27.66) с точностью до знака совпадает с $R_{1(2)}$.

Отсеивание переменных в исследовательской работе

27.27 В новых областях исследования предварительное изучение связей между величинами часто начинается с вычисления всевозможных корреляций нулевого порядка, образующих корреляционную матрицу $\boldsymbol{C}$. Если нас интересует лишь «предсказание» значения одной величины, x_1, по другим, то заманчиво было бы вычислить вначале только корреляции между x_1 и остальными величинами и отбросить те величины, с которыми корреляции равны нулю или очень малы. Это, возможно, позволило бы уменьшить число величин до поддающегося обработке количества. Следующий этап состоял бы в вычислении корреляционной матрицы оставшихся величин и множественных корреляций между x_1 и комбинациями из остальных величин.

К сожалению, эта процедура может привести к серьезному заблуждению. Совершенно верно, что множество корреляций нулевого порядка полностью определяет всю совокупность частных корреляций. Но малость коэффициентов нулевого порядка между x_1 и другими величинами не гарантирует малости коэффициентов более высокого порядка, и это так, даже если пренебрегать выборочными рассмотрениями. Поскольку согласно (27.62) множественная корреляция должна быть не меньше наибольшей из корреляций любого порядка, то с помощью описанной выше процедуры «отсеивания» мы, может быть, отбрасываем ценную информацию.

Рассмотрим снова (27.5):

$$\rho_{12.3} = \frac{\rho_{12} - \rho_{13}\rho_{23}}{\{(1-\rho_{13}^2)(1-\rho_{23}^2)\}^{1/2}}. \qquad (27.67)$$

Если ρ_{12} и ρ_{13} либо ρ_{23} равны нулю, то и $\rho_{12.3}=0$; если $\rho_{13}=\rho_{23}=0$, то $\rho_{12.3}=\rho_{12}$. *Но* ρ_{12} *и* ρ_{13} *могут оба быть очень малыми, тогда как* $\rho_{12.3}$ *может быть очень велик.* Действительно, предположим, что $\rho_{13}=0$. Тогда (27.67) превращается в

$$\rho_{12.3} = \rho_{12}/(1-\rho_{23}^2)^{1/2}, \qquad \rho_{13}=0. \qquad (27.68)$$

Если ρ_{12} очень мал, а ρ_{23}^2 очень велик, то (27.68) может быть также очень большим. Рассмотрим конкретный пример. Пусть

$$\rho_{13}=0, \qquad \rho_{12}=\cos\theta, \qquad \rho_{23}=\cos\left(\tfrac{1}{2}\pi-\theta\right)=\sin\theta.$$

Тогда (27.68) имеет вид $\rho_{12.3}=1$. Аналогичный результат получится, если положить $\rho_{23}=\cos\left(\frac{1}{2}\pi+\theta\right)$, так как при этом ρ_{23}^2 не изменится.

Теперь можно сделать $\cos\theta$ (или $\cos\left(\frac{1}{2}\pi+\theta\right)$) как угодно малым, скажем равным ε. Тогда

$$\rho_{13}=0, \quad \rho_{12}=\varepsilon, \quad \rho_{12.3}=1. \tag{27.69}$$

Поскольку согласно (27.62) множественная корреляция $R_{1(23)} \geqslant |\rho_{12.3}|$, то в этом случае находим, что

$$R_{1\,(23)}=1. \tag{27.70}$$

В силу (27.70) x_1 является точной линейной функцией от x_2 и x_3, вопреки значениям коэффициентов нулевого порядка в (27.69). Очевидно, было бы неразумно на основании только коэффициентов нулевого порядка отбрасывать x_2 и x_3 в задаче предсказания x_1.

В геометрических терминах легко понять, что здесь произошло. Вектор PQ_1 ортогонален к PQ_3 и почти ортогонален (при угле θ, близком к $\frac{1}{2}\pi$) к PQ_2. Но все три вектора лежат в одной и той же плоскости, в которой PQ_2 и PQ_3 образуют между собой либо угол $\left(\frac{1}{2}\pi-\theta\right)$ (тогда $\cos\left(\frac{1}{2}\pi-\theta\right)$ близок к 1), либо угол $\left(\frac{1}{2}\pi+\theta\right)$ (тогда $\cos\left(\frac{1}{2}\pi+\theta\right)$ близок к -1).

Мы рассмотрели простой пример, но то же самое рассуждение *a fortiori* применимо к большему числу переменных, где имеется больше места для появления связей указанного сорта. Значение R зависит от *всех* частных корреляций.

К счастью для человеческого нетерпения, жизнь обычно оказывается менее сложной, чем она могла бы быть, и мы обычно избегаем наихудших возможных последствий упрощенного выбора независимых переменных. Обыкновенно, даже в новых областях, у нас имеется достаточно предварительных знаний, помогающих избежать наиболее вопиющих оплошностей, но логическая трудность остается.

Выборочный множественный коэффициент корреляции и его условное распределение

27.28 Определим теперь выборочный аналог для $R^2_{1\,(2\ldots p)}$, полагая

$$1-R^2_{1\,(2\ldots p)}=\frac{s^2_{1.2\ldots p}}{s^2_1}. \tag{27.71}$$

Все соотношения из **27.23—26** остаются в силе с заменой ρ на r и σ на s. Мы переходим к детальному обсуждению выборочного распределения величины R^2. Поскольку согласно **27.23** это есть

коэффициент корреляции, значение которого не зависит от расположения и масштаба, то его распределение не будет зависеть от параметров расположения и масштаба.

Вначале рассмотрим условное распределение величины R^2, когда значения $x_2, \ldots, x_p$ фиксированы. Как и в (26.50), учитывая (27.71), запишем тождество

$$ns_1^2 \equiv ns_1^2 R^2_{1\,(2\ldots p)} + ns_1^2\left(1 - R^2_{1\,(2\ldots p)}\right) \equiv$$
$$\equiv n\left(s_1^2 - s^2_{1.2\ldots p}\right) + ns^2_{1.2\ldots p}. \quad (27.72)$$

Если наблюдения над x_1 являются независимыми нормальными величинами, так что $\boldsymbol{R}^2_{1\,(2\ldots p)} = 0$, то левая часть в (27.72) подчиняется хи-квадрат распределению с $n-1$ степенями свободы. Квадратичные формы относительно x_1 в правой части (27.72), как можно показать, имеют ранги $p-1$ и $n-p$ соответственно. Из теоремы Кокрэна (**15.16**) следует, что они независимы и подчинены хи-квадрат распределениям с теми же числами степеней свободы и что отношение

$$F = \frac{R^2_{1\,(2\ldots p)}/(p-1)}{\left(1 - R^2_{1\,(2\ldots p)}\right)/(n-p)} \quad (27.73)$$

имеет F-распределение с $(p-1,\ n-p)$ степенями свободы. Приведенный результат впервые был доказан Фишером (1924b). Частным случаем (27.73) при $p=2$, когда $R^2_{1\,(2)} = r^2_{12}$ (см. (27.66)), является (26.52).

Это еще один пример критерия ОП для линейной гипотезы. Мы постулируем, что среднее значение наблюдений над x_1 является линейной функцией от $p-1$ других величин с $p-1$ коэффициентами плюс постоянный член, так что всего имеется p параметров. Проверяется гипотеза о том, что все $p-1$ коэффициентов равны нулю, т. е. $H_0 : \boldsymbol{R}^2 = 0$. В обозначениях пунктов **24.27—28** $k=p$, $r=p-1$. Таким образом, как мы и видели, F-критерий (27.73) имеет $(p-1,\ n-p)$ степеней свободы. Из **24.32—33** немедленно следует, что если гипотеза H_0 неверна, то величина F в (27.73) подчинена нецентральному F-распределению с $(p-1,\ n-p)$ степенями свободы и параметром нецентральности $\lambda = n\boldsymbol{R}^2$, и свойства мощности критерия ОП, приведенные в главе 24, приложимы здесь. В частности, согласно **24.36—37** этот критерий является РНМ инвариантным критерием.

Многомерный нормальный (безусловный) случай

27.29 Пусть теперь значения $x_2, \ldots, x_p$ также изменяются, и предположим, что выборка производится из многомерной нормальной совокупности. Оказывается, что распределение величины R^2 остается прежним при $\boldsymbol{R}^2 = 0$, но при $\boldsymbol{R}^2 \neq 0$ совер-

шенно отличается от распределения R^2 с фиксированными $x_2, \ldots, x_p$. Таким образом, функции мощности критерия для $\boldsymbol{R}^2 = 0$ различны в этих двух случаях, хотя один и тот же критерий имеет силу в каждом случае. Однако, когда $n \to \infty$, оба результата совпадают.

Мы выведем геометрически результат для многомерного нормального случая с $\boldsymbol{R}^2 = 0$, а в **27.30** обобщим его.

Рассмотрим геометрическое представление из **27.26**. Тогда R есть косинус угла, скажем θ, между PQ_1 (вектор наблюдений x_1) и вектором PV, лежащим в $(p-1)$-мерном пространстве S_{p-1} остальных переменных и минимизирующим угол с PQ_1. Если теоретический коэффициент $\boldsymbol{R} = 0$, то вследствие многомерной нормальности совокупности x_1 не зависит от $x_2, \ldots, x_p$, а тогда вектор PQ_1, из-за радиальной симметрии нормального распределения, будет иметь случайное направление относительно подпространства S_{p-1}, которое в последующих рассуждениях можно считать фиксированным. (Этим объясняется, почему совпадают условные и безусловные результаты при $\boldsymbol{R}^2 = 0$).

Мы должны рассмотреть относительные вероятности, с которыми могут появляться различные значения угла θ. При фиксированной дисперсии s_1^2 функция плотности выборки из n наблюдений постоянна на $(n-2)$-мерной поверхности $(n-1)$-мерной гиперсферы. Если θ и PV фиксированы, то вектор PQ_1 будет лежать на гиперсфере размерности $(n-2)-(p-1) = (n-p-1)$, объем которой пропорционален $(\sin\theta)^{n-p-1}$ (см. **16.24**).

Рассмотрим теперь, что происходит, если вектор PV изменяется. PV может изменяться внутри S_{p-1}, где вследствие радиальной симметрии его значения равновероятны на $(p-2)$-мерной поверхности сферы размерности $(p-1)$. Эта поверхность имеет объем, пропорциональный $(\cos\theta)^{p-2}$. Следовательно, для фиксированного θ получаем элемент вероятности $(\sin\theta)^{n-p-1}(\cos\theta)^{p-2}d\theta$. Полагая $R = \cos\theta$ и

$$d\theta \propto d(R^2)/\{R(1-R^2)^{1/2}\},$$

находим, что R^2 имеет бета-распределение

$$dF \propto (R^2)^{(p-3)/2}(1-R^2)^{(n-p-2)/2}\,d(R^2), \quad 0 \leqslant R^2 \leqslant 1. \qquad (27.74)$$

Легко видеть, что нормирующая постоянная равна $\dfrac{1}{\mathrm{B}\{(p-1)/2,\ (n-p)/2\}}$. Применяя преобразование (27.73) к (27.74), получаем в точности то же F-распределение, которое

было выведено в **27.28** при фиксированных $x_2, \ldots, x_p$. Когда $p = 2$, (27.74) сводится к (16.62), с той лишь разницей, что там используется dR, а не $d(R^2)$.

27.30 Перейдем теперь к случаю, когда $\boldsymbol{R} \neq 0$. Распределение коэффициента R для этого случая было впервые найдено Фишером (1928а) с помощью значительного развития геометрического подхода из **27.29**. Мы приведем более простой вывод, принадлежащий Морэну (1950).

Соотношение (27.61) для выборочного коэффициента можно записать в виде

$$1 - R^2_{1(2\ldots p)} = (1 - r^2_{12})(1 - T^2), \tag{27.75}$$

где T есть множественный коэффициент корреляции между $x_{1.2}$ и $x_{3.2}, x_{4.2}, \ldots, x_{p.2}$. Коэффициент $\boldsymbol{R}_{1(2\ldots p)}$ и распределение величины $R_{1(2\ldots p)}$ не изменятся, если $x_2, \ldots, x_p$ подвергнуть ортогональному преобразованию. При этом преобразовании в качестве новой переменной x_2 выберем такую линейную функцию от старых переменных $x_2, \ldots, x_p$, которая имеет с x_1 максимальную корреляцию, т. е. $\rho_{12} = \boldsymbol{R}_{1(2\ldots p)}$. Тогда из (27.61) вытекает, что

$$\rho^2_{13.2} = \rho^2_{14.23} = \ldots = \rho^2_{1p.23\ldots(p-1)} = 0, \tag{27.76}$$

а поскольку в (27.61) можно переставлять индексы, отличные от 1, то *все* частные коэффициенты вида $\rho_{1j.2s}$ равны нулю. Таким образом, величина $x_{1.2}$ некоррелирована с (а поскольку распределение нормально, то и не зависит от) $x_{3.2}, x_{4.2}, \ldots, x_{p.2}$, и коэффициент T в (27.75) распределен, как множественный коэффициент корреляции между одной величиной и $p - 2$ другими, основанный на $n - 1$ наблюдениях (одно измерение теряется при проектировании для получения остатков), при теоретическом $\boldsymbol{R} = 0$. Кроме того, коэффициент T распределен независимо от r_{12}, потому что все величины $x_{j.2}$, согласно (27.46), ортогональны к x_2. Итак, множители в правой части (27.75) независимы. Распределение коэффициента r_{12}, скажем $f_1(r)$, получается из (16.60) с $\rho = \boldsymbol{R}_{1(2\ldots p)}$ интегрированием относительно β по области его изменения. Распределение же величины T^2, скажем $f_2(R^2)$, имеет вид (27.74) с n и p, уменьшенными на 1. Следовательно, из (27.75) находим распределение коэффициента R^2:

$$dF = \int_{-R}^{R} \left\{ f_2\left(\frac{R^2 - r^2}{1 - r^2}\right) \right\} dF_1(r). \tag{27.77}$$

Опуская для удобства все индексы, получаем

$$dF = \frac{(n-2)}{\pi}(1-\boldsymbol{R}^2)^{(n-1)/2}\int_{-R}^{R}(1-r^2)^{(n-4)/2}\int_0^{\infty}\frac{d\beta}{(\mathrm{ch}\beta-\boldsymbol{R}r)^{n-1}}\times$$

$$\times\left\{\frac{1}{\mathrm{B}\{(p-2)/2,\,(n-p)/2\}}\left(\frac{R^2-r^2}{1-r^2}\right)^{(p-4)/2}\left(\frac{1-R^2}{1-r^2}\right)^{(n-p-2)/2}\frac{dR^2}{(1-r^2)}\right\}dr=$$

$$=\frac{(n-2)}{\pi}\cdot\frac{(1-\boldsymbol{R}^2)^{(n-1)/2}(1-R^2)^{(n-p-2)/2}\,d\,(R^2)}{\mathrm{B}\{(p-2)/2,\,(n-p)/2\}}\int_{-R}^{R}(R^2-r^2)^{(p-4)/2}\times$$

$$\times\left[\int_0^{\infty}\frac{d\beta}{(\mathrm{ch}\beta-\boldsymbol{R}r)^{n-1}}\right]dr. \quad (27.78)$$

Полагая в (27.78) $r = R\cos\psi$ и беря интеграл по β от $-\infty$ до ∞ (для компенсации последнего интеграл надо разделить на 2), получим распределение в форме Фишера:

$$dF=\frac{\Gamma(n/2)\,(1-\boldsymbol{R}^2)^{(n-1)/2}}{\pi\Gamma\{(p-2)/2\}\,\Gamma\{(n-p)/2\}}(R^2)^{(p-3)/2}(1-R^2)^{(n-p-2)/2}\,d\,(R^2)\times$$

$$\times\int_0^{\pi}\sin^{p-3}\psi\left\{\int_{-\infty}^{\infty}\frac{d\beta}{(\mathrm{ch}\,\beta-\boldsymbol{R}R\cos\psi)^{n-1}}\right\}d\psi. \quad (27.79)$$

27.31 Распределение (27.79) может быть выражено через гипергеометрическую функцию. Раскладывая подынтегральную функцию в равномерно сходящийся ряд по степеням $\cos\psi$ и отбрасывая нечетные степени $\cos\psi$, так как при интегрировании от 0 до π они исчезнут, получаем

$$\sum_{j=0}^{\infty}\binom{n+2j-2}{2j}\frac{\sin^{p-3}\psi\cos^{2j}\psi}{(\mathrm{ch}\beta)^{n-1+2j}}(\boldsymbol{R}R)^{2j}.$$

Поскольку

$$\int_0^{\pi}\cos^{2j}\psi\sin^{p-3}\psi\,d\psi=\mathrm{B}\left\{\tfrac{1}{2}(p-2),\,\tfrac{1}{2}(2j+1)\right\}$$

и

$$\int_{-\infty}^{\infty}\frac{d\beta}{(\mathrm{ch}\beta)^{n-1+2j}}=\mathrm{B}\left\{\frac{1}{2},\,\frac{1}{2}(n+2j-1)\right\},$$

то интеграл в (27.79) превращается в

$$\sum_{j=0}^{\infty}\binom{n+2j-2}{2j}\mathrm{B}\left\{\tfrac{1}{2}(p-2),\,\tfrac{1}{2}(2j+1)\right\}\mathrm{B}\left\{\frac{1}{2},\,\frac{1}{2}(n+2j-1)\right\}(\boldsymbol{R}R)^{2j},$$

что после применения гамма-функции и упрощений принимает вид

$$\frac{\pi\Gamma\{(p-2)/2\}\,\Gamma\{(n-1)/2\}}{\Gamma(n/2)\,\Gamma\{(p-1)/2\}}\,F\left\{\tfrac{1}{2}(n-1),\ \tfrac{1}{2}(n-1),\ \tfrac{1}{2}(p-1),\ \boldsymbol{R}^2R^2\right\}. \tag{27.80}$$

Подставляя (27.80) в качестве подинтегральной функции в (27.79), получаем

$$dF=\frac{(R^2)^{(p-3)/2}(1-R^2)^{(n-p-2)/2}\,d(R^2)}{\mathrm{B}\{(p-1)/2,\ (n-p)/2\}}(1-\boldsymbol{R}^2)^{(n-1)/2}\times$$
$$\times F\left\{\tfrac{1}{2}(n-1),\ \tfrac{1}{2}(n-1),\ \tfrac{1}{2}(p-1),\ \boldsymbol{R}^2R^2\right\}. \tag{27.81}$$

Это безусловное распределение следует сравнить с условным распределением величины R^2, приведенным в упражнении 27.13 и легко получаемым из нецентрального F-распределения, данного в **27.28**. Упражнение 27.14 показывает, что при $n\to\infty$ оба распределения дают для nR^2 нецентральное χ^2-распределение.

Первый множитель в правой части (27.81) есть распределение (27.74) для $\boldsymbol{R}=0$; в этом случае второй множитель превращается в единицу. Если $p=2$, то (27.81) является не столь быстро сходящимся рядом по r^2, как (16.66) по r. И вообще он сходится медленно, потому что первые два аргумента в гипергеометрической функции равны $(n-1)/2$. В поисках более быстро сходящегося выражения соблазнительно в (27.78) вместо интеграла по β подставить выражение (16.65), которое имеет вид

$$\int_0^\infty\frac{d\beta}{(\mathrm{ch}\beta-\boldsymbol{R}r)^{n-1}}=\frac{\mathrm{B}(1/2,\ n-1)}{2^{1/2}(1-\boldsymbol{R}r)^{n-3/2}}F\left\{\tfrac{1}{2},\ \tfrac{1}{2},\ n-\tfrac{1}{2},\ \tfrac{1}{2}(1+\boldsymbol{R}r)\right\},$$

а так как $F(a,b,c,x)=(1-x)^{c-a-b}F(c-a,c-b,c,x)$, то оно равно

$$\frac{\mathrm{B}(1/2,\ n-1)}{2^{n-1}}F\left\{n-1,\ n-1,\ n-\tfrac{1}{2},\ \tfrac{1}{2}(1+\boldsymbol{R}r)\right\}. \tag{27.82}$$

Но когда мы подставим (27.82) в (27.78), то интегрирование по r вряд ли приведет к более удобному результату, чем (27.81).

Моменты и предельные распределения величины R^2

27.32 Можно показать (см. Уишарт (1931)), что среднее значение величины R^2 в многомерном нормальном случае равно

$$\mathbf{M}(R^2)=1-\frac{n-p}{n-1}(1-\boldsymbol{R}^2)\,F\left\{1,\ 1,\ \tfrac{1}{2}(n+1),\ \boldsymbol{R}^2\right\}=$$
$$=\boldsymbol{R}^2+\frac{p-1}{n-1}(1-\boldsymbol{R}^2)-\frac{2(n-p)}{n^2-1}\boldsymbol{R}^2(1-\boldsymbol{R}^2)+O\left(\frac{1}{n^2}\right). \tag{27.83}$$

В частности, когда $\boldsymbol{R}^2 = 0$, (27.83) сводится к

$$\mathsf{M}(R^2 \mid \boldsymbol{R}^2 = 0) = \frac{p-1}{n-1}, \tag{27.84}$$

что также можно получить непосредственно из (27.74).

Аналогично может быть показано, что дисперсия равна

$$\mathsf{D}(R^2) = \frac{(n-p)(n-p+2)}{(n^2-1)}(1-\boldsymbol{R}^2)^2 F\left(2, 2, \tfrac{1}{2}(n+3), \boldsymbol{R}^2\right) - \\ - \{\mathsf{M}(R^2) - 1\}^2 = \tag{27.85}$$

$$= \frac{(n-p)}{(n^2-1)(n-1)}(1-\boldsymbol{R}^2)^2\left[2(p-1) + \frac{4\boldsymbol{R}^2\{(n-p)(n-1)+4(p-1)\}}{(n+3)} + \right. \\ \left. + O\left\{\left(\frac{\boldsymbol{R}^2}{n}\right)^2\right\}\right]. \tag{27.86}$$

Соотношение (27.86) представимо в виде

$$\mathsf{D}(R^2) = \frac{4\boldsymbol{R}^2(1-\boldsymbol{R}^2)^2(n-p)^2}{(n^2-1)(n+3)} + O\left(\frac{1}{n^2}\right), \tag{27.87}$$

так что если $\boldsymbol{R}^2 \neq 0$, то

$$\mathsf{D}(R^2) \sim 4\boldsymbol{R}^2(1-\boldsymbol{R}^2)^2/n. \tag{27.88}$$

Но если $\boldsymbol{R}^2 = 0$, то равенство (27.87) оказывается бесполезным. Возвращаясь же к (27.86), находим

$$\mathsf{D}(R^2) = \frac{2(n-p)(p-1)}{(n^2-1)(n-1)} \sim \frac{2(p-1)}{n^2}. \tag{27.89}$$

Точный результат в (27.89) можно было получить из (27.74).

27.33 Различные порядки величины асимптотических дисперсий (27.88) и (27.89), когда $\boldsymbol{R} \neq 0$ и $\boldsymbol{R} = 0$, отражают существенно разное поведение распределения коэффициента R^2 в этих двух ситуациях. Хотя (27.84) показывает, что R^2 является смещенной оценкой для $\boldsymbol{R}^2$, тем не менее она, очевидно, состоятельна. Для больших n $\mathsf{M}(R^2) \to \boldsymbol{R}^2$, а $\mathsf{D}(R^2) \to 0$. При $\boldsymbol{R} \neq 0$ распределение величины R^2 асимптотически нормально со средним $\boldsymbol{R}^2$ и дисперсией (27.88) (см. упражнение 27.15). Однако, когда $\boldsymbol{R} = 0$, коэффициент R, который заключен в интервале $(0, 1)$, сходится к 0, нижней границе своего интервала изменения, и одного этого достаточно, чтобы показать, что в данном случае его распределение не будет нормальным (см. упражнения 27.14, 27.15). При таких обстоятельствах неудивительно, что его дисперсия имеет порядок n^{-2}: аналогичная ситуация возникала при оценивании границы конечного промежутка, на котором сосредоточено распределение, где, как мы видели в упражнениях 14.8, 14.13, 14.16, появляются дисперсии порядка n^{-2}.

Распределение коэффициента R ведет себя аналогично (в отношении предельной нормальности) распределению R^2, хотя мы увидим, что его дисперсия имеет всегда порядок $1/n$.

Следует упомянуть одно прямое следствие особенности в распределении коэффициента R при $\boldsymbol{R}^2 = 0$. Из (27.88) следует, что

$$\mathsf{D}R \sim (1 - \boldsymbol{R}^2)^2/n; \tag{27.90}$$

это совпадает с асимптотическим выражением для дисперсии обычного коэффициента корреляции (см. (26.24))

$$\mathsf{D}r \sim (1 - \rho^2)^2/n.$$

Естественно применить к R стабилизирующее дисперсию z-преобразование из **16.33** (см. также упражнение 16.18), дающее величину $z = \operatorname{arg\,th} R$, дисперсия которой близка к $1/n$ независимо от значения $\boldsymbol{R}$. Но, как отметил Хотеллинг (1953), это неверно вблизи $\boldsymbol{R} = 0$, поскольку там нарушается (27.90). В этом случае асимптотическая дисперсия в силу соотношения (27.84) равна

$$\mathsf{D}R = \mathsf{M}(R^2) - \{\mathsf{M}(R)\}^2 \sim (p - 1)/n, \tag{27.91}$$

что отличается от значения $1/n$, получаемого из (27.90). Все хорошо при $p = 2$ (когда $R = |r|$). В остальных случаях z-преобразование величины R можно использовать только для значений $\boldsymbol{R}$, отделенных от нуля.

Несмещенное оценивание коэффициента R^2 в многомерном нормальном случае

27.34 Поскольку согласно (27.83) R^2 представляет собой смещенную оценку для $\boldsymbol{R}^2$, то мы можем попытаться устранить смещение. Олкин и Прэтт (1958) показали, что несмещенной оценкой коэффициента $\boldsymbol{R}^2_{1\,(2\,\ldots\,p)}$ служит

$$t = 1 - \frac{n-3}{n-p}\left(1 - R^2_{1\,(2\,\ldots\,p)}\right) F\left(1,\ 1,\ \tfrac{1}{2}(n - p + 2),\ 1 - R^2_{1\,(2\,\ldots\,p)}\right), \tag{27.92}$$

где $n > p \geqslant 3$. Статистика t есть единственная несмещенная функция от R^2, так как она является функцией от полной системы достаточных статистик. Правую часть (27.92) можно разложить в ряд

$$t = R^2 - \frac{p-3}{n-p}(1 - R^2) - \left\{\frac{2(n-3)}{(n-p)(n-p+2)}(1 - R^2)^2 + O\left(\frac{1}{n^2}\right)\right\}, \tag{27.93}$$

откуда следует, что $t \leqslant R^2$. Если $R^2 = 1$, то и $t = 1$. С другой стороны, когда R^2 мало или равно нулю, то t отрицательно, как этого можно было ожидать. Нельзя найти несмещенную оценку

для R^2 (т. е. оценку, математическое ожидание которой равно R^2 *при любом истинном значении* R^2), которая бы принимала только неотрицательные значения, хотя мы знаем, что коэффициент R^2 неотрицателен. Можно устранить абсурдность отрицательных значений оценки, используя в качестве новой оценки

$$t' = \max(t, 0), \tag{27.94}$$

но (27.94) перестает быть несмещенной оценкой.

27.35 Леман (1959) показал, что критерии для проверки гипотезы $R^2 = 0$, отвергающие гипотезу при больших значениях R^2, являются РНМ критериями в классе критериев со статистиками, инвариантными относительно изменений масштабов и сдвигов. Эзекиел и Фокс (1959), а также К. Крамер (1963) приводят графики и таблицы для построения доверительных интервалов для R^2 по значениям R^2.

УПРАЖНЕНИЯ

27.1 Показать, что

$$\beta_{12.34\ldots(p-1)} = \frac{\beta_{12.34\ldots p} + \beta_{1p.23\ldots(p-1)}\beta_{p2.13\ldots(p-1)}}{1 - \beta_{1p.23\ldots(p-1)}\beta_{p1.23\ldots(p-1)}}$$

и

$$\rho_{12.34\ldots(p-1)} = \frac{\rho_{12.34\ldots p} + \rho_{1p.23\ldots(p-1)}\rho_{2p.13\ldots(p-1)}}{\left\{\left(1 - \rho^2_{1p.23\ldots(p-1)}\right)\left(1 - \rho^2_{2p.13\ldots(p-1)}\right)\right\}^{1/2}}.$$

(Юл, 1907.)

27.2 Показать, что для p случайных величин имеется $\binom{p}{2}$ коэффициентов корреляции нулевого порядка и $\binom{p-2}{s}\binom{p}{2}$ порядка s, а всего имеется $\binom{p}{2}2^{p-2}$ коэффициентов корреляции и $\binom{p}{2}2^{p-1}$ коэффициентов регрессии.

27.3 Пусть для некоторого множества величин все коэффициенты корреляции нулевого порядка равны ρ. Показать, что тогда каждый частный коэффициент корреляции s-го порядка равен $\dfrac{\rho}{1 + s\rho}$.

27.4 Доказать равенство (27.27). Используя его, показать, что коэффициент при $x_l x_m$ в экспоненте многомерного нормального распределения случайных величин $x_1, x_2, \ldots, x_p$ равен $1/\mathrm{cov}\left(x_l \cdot q_l,\ x_m \cdot q_m\right)$.

27.5 Исходя из (27.46), показать, что при суммировании попарных произведений остатков можно опускать любой или все вторичные индексы у того остатка, *все* вторичные индексы которого содержатся среди вторичных индексов второго сомножителя, т. е. что

$$\sum x_{1.stu}x_{2.st} = \sum x_{1.stu}x_{2.s} = \sum x_{1.stu}x_{2},$$

но что

$$\sum x_{1.stu}x_{2.st} \neq \sum x_{1.su}x_{2.st},$$

где s, t, u — множества индексов.

(Чандлер, 1950.)

27.6 С помощью преобразования

$$y_1 = x_1,$$
$$y_2 = x_{2.1},$$
$$y_3 = x_{3.21},$$
$$\ldots\ldots$$

доказать, что многомерное нормальное распределение может быть записано в виде

$$dF = \frac{1}{(2\pi)^{p/2}\sigma_1\sigma_{2.1}\sigma_{3.12}\ldots}\exp\left\{-\frac{1}{2}\left(\frac{x_1^2}{\sigma_1^2}+\frac{x_{2.1}^2}{\sigma_{2.1}^2}+\frac{x_{3.12}^2}{\sigma_{3.12}^2}+\ldots\right)\right\}dx_1\,dx_{2.1}\ldots,$$

так что остатки x_1, $x_{2.1}$ ... не зависят один от другого. Вывести отсюда, что любые два остатка $x_{j.r}$ и $x_{k.r}$ (где r — множество общих индексов) имеют двумерное нормальное распределение с коэффициентом корреляции $\rho_{jk.r}$.

27.7 Показать, что если применить ортогональное преобразование к множеству из n независимых наблюдений над p-мерной нормальной величиной, то преобразованные n наблюдений также будут независимыми.

27.8 Для данных из таблиц 26.1 и 26.2 мы видели в примере 26.6 и упражнении 26.1, что $r_{12} = 0{,}34$, $r_{13} = 0{,}07$, где индексы 1, 2, 3 относятся к росту, весу и объему груди соответственно. Пусть также дано, что $r_{23} = 0{,}86$. Установить равенство

$$R^2_{3\,(12)} = 0{,}80,$$

указывающее на то, что объем груди достаточно хорошо определяется линейной функцией от роста и веса.

27.9 Показать непосредственно, что никакая линейная функция от $x_2, \ldots, x_p$ не имеет более высокой корреляции с x_1, чем оценка наименьших квадратов величины x_1.

27.10 Установить выражение (27.83) для $\mathbf{M}(R^2)$.

(Уишарт, 1931.)

27.11 Установить выражение (27.85) для $\mathbf{D}(R^2)$.

(Уишарт, 1931.)

27.12 Проверить, что (27.92) является несмещенной оценкой коэффициента R^2.

27.13 Пусть $R^2 \neq 0$. Исходя из нецентрального F-распределения величины F в (27.73), показать, что в этом случае коэффициент R^2 при фиксированных $x_2, \ldots, x_p$ имеет распределение

$$dF = \frac{1}{\mathrm{B}\{(p-1)/2,\,(n-p)/2\}}(R^2)^{(p-3)/2}(1-R^2)^{(n-p-2)/2}\,dR^2\exp\left\{-\frac{1}{2}(n-p)R^2\right\}\times$$
$$\times\sum_{j=0}^{\infty}\frac{\Gamma\{(n-1+2j)/2\}\,\Gamma\{(p-1)/2\}}{\Gamma\{(n-1)/2\}\,\Gamma\{(p-1+2j)/2\}}\,\frac{\{(1/2)(n-p)R^2R^2\}^j}{j!}.$$

(Фишер, 1928а.)

27.14 Используя (27.81), доказать, что при $n \to \infty$ и фиксированном p величина $nR^2 = B^2$ имеет распределение

$$dF = \frac{(B^2)^{(p-3)/2}}{2^{(p-1)/2}\Gamma\{(p-1)/2\}} \exp\left(-\frac{1}{2}\beta^2 - \frac{1}{2}B^2\right) \times \left\{1 + \frac{\beta^2 B^2}{(p-1)\cdot 2} + \frac{(\beta^2 B^2)^2}{(p-1)(p+1)\cdot 2\cdot 4} + \ldots\right\} d(B^2),$$

где $\beta^2 = n\boldsymbol{R}^2$, и что, следовательно, nR^2 является нецентральной χ^2-величиной вида (24.18) с $\nu = p - 1$, $\lambda = n\boldsymbol{R}^2$. На основании упражнения 27.13 показать, что этот же результат верен для условного распределения величины nR^2.

(Фишер, 1928а.)

27.15 В условиях упражнения 27.14 с помощью х. ф. нецентрального χ^2-распределения (см. упражнение 24.1) доказать, что если $n \to \infty$, а p фиксировано, то коэффициент R^2 имеет асимптотически нормальное распределение, когда $\boldsymbol{R} \neq 0$, и что это неверно, когда $\boldsymbol{R} = 0$. Распространить результат на коэффициент R.

27.16 Пусть коэффициент R^2 вычислен по выборке из многомерного распределения. Показать, что если $n - p$ четно, то его функция распределения может быть записана в виде

$$(1 - \boldsymbol{R}^2)^{(n-1)/2} R^{p-1} \sum_{j=0}^{\frac{1}{2}(n-p-2)} \frac{\Gamma\{(p-1+2j)/2\}}{\Gamma\{(p-1)/2\}} \cdot \frac{(1-R^2)^j}{(1-\boldsymbol{R}^2R^2)^{(n-1+2j)/2}} \times F\left(-j, -\frac{1}{2}(n-p), \frac{1}{2}(p-1), \boldsymbol{R}^2R^2\right).$$

(Фишер, 1928а.)

27.17 Дана выборка $(x_1, \ldots, x_n)$, состоящая из одного наблюдения над n-мерной нормальной случайной величиной, все средние значения которой равны μ, все дисперсии σ^2, а все коэффициенты корреляции ρ. Доказать, что статистика

$$t^2 = \frac{(\bar{x} - \mu)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2/\{n(n-1)\}} \cdot \left\{\frac{1-\rho}{1+(n-1)\rho}\right\}$$

имеет t-распределение Стьюдента с $n - 1$ степенями свободы. Когда $\rho = 0$, это сводится к обычному критерию о среднем значении для n независимых нормальных величин.

(Уолш, 1947.)

27.18 Пусть $x_0, x_1, \ldots, x_n$ — нормальные величины с одной и той же дисперсией α^2, $x_1, \ldots, x_n$ независимы между собой, а x_0 имеет среднее нуль и корреляцию λ со всеми остальными величинами. Показать, что n величин

$$y_i = x_i - ax_0, \qquad i = 1, 2, \ldots, n,$$

имеют многомерное нормальное распределение, в котором все корреляции равны

$$\rho = (a^2 - 2a\lambda)/(1 + a^2 - 2a\lambda)$$

и все дисперсии равны

$$\sigma^2 = \alpha^2/(1-\rho).$$

(Стьюарт, 1958.)

27.19 Используя результат упражнения 27.18, получить утверждение упражнения 27.17.

(Стьюарт, 1958.)

27.20 Доказать, что если каждая пара из $x_2, \ldots, x_p$ некоррелирована, то

$$R^2_{1(2\ldots p)} = \sum_{s=2}^{p} \rho^2_{1s} = \sum_{s=2}^{p} (\beta_{1s.q}\sigma_s)^2/\sigma_1^2.$$

27.21 Обобщая (27.17) и используя матричные обозначения, данные в конце **27.6**, показать, что условное среднее вектора $(x_1, \ldots, x_h)'$ при фиксированном значении $\boldsymbol{x}_0'$ вектора $(x_{h+1}, \ldots, x_p)'$ равно $\boldsymbol{B}'\boldsymbol{E}^{-1}\boldsymbol{x}_0'$.

(Марсалья (1964) показал, что этот результат и (27.14) верны даже для вырожденного многомерного нормального распределения, если только $\boldsymbol{E}^{-1}$ заменить на *псевдообратную* матрицу $\boldsymbol{E}^{+} = \boldsymbol{T}'(\boldsymbol{T}\boldsymbol{T}')^{-2}\boldsymbol{T}$, где $\boldsymbol{E} = \boldsymbol{T}'\boldsymbol{T}$.)

ГЛАВА 28

ОБЩАЯ ТЕОРИЯ РЕГРЕССИИ

28.1 В последних двух главах мы изложили теорию линейной регрессии одной величины относительно одной или большего числа других величин, но основное внимание там уделялось теории корреляции. Теперь же центром нашего исследования станет теория регрессии. В этой главе мы обобщим и объединим результаты глав 26 и 27, при этом будет использоваться метод наименьших квадратов, изложенный в главе 19.

При рассмотрении регрессии y относительно одной или большего числа величин x обычно называют y «зависимой», а x «независимыми» переменными. Эта терминология, заимствованная из обычной алгебры, плоха, потому что величины x не являются, в общем случае, независимыми в вероятностном смысле. В действительности мы увидим, что они вовсе не обязательно являются случайными величинами. Далее, поскольку цель регрессионного анализа состоит в исследовании зависимости y от x, то название «независимые» для величин x вносит особенную путаницу. Поэтому, несмотря на общепринятую терминологию, мы будем следовать некоторым более современным авторам, например Э. Хеннану (1956), и называть x *регрессионными переменными* (или, для краткости, *регрессорами*).

Вначале мы рассмотрим обобщение аналитической теории регрессии, изучавшейся в главах 26 и 27 для линейного случая. Отличительная особенность аналитической теории состоит в том, что предполагается известным совместное распределение величин или, что эквивалентно, их совместная характеристическая функция.

Аналитическая теория регрессии

28.2 Пусть $f(x, y)$ есть совместная плотность распределения величин x, y. Тогда для любого фиксированного значения x, скажем X, математическое ожидание величины y определяется как

$$\mathbf{M}(y \mid X) = \int_{-\infty}^{\infty} yf(X, y)\,dy \Big/ \int_{-\infty}^{\infty} f(X, y)\,dy. \qquad (28.1)$$

Соотношение (28.1) определяет *регрессию* (*кривую регрессии*), обсуждавшуюся в **26.5**. Оно выражает зависимость между X и средним значением величины y для этого значения X; это зависимость математическая, а не вероятностная.

Можно также рассмотреть более общую регрессию (кривую регрессии) порядка r, определяемую равенством

$$\mu'_{rX} = \mathbf{M}(y^r \mid X) = \int_{-\infty}^{\infty} y^r f(X, y)\, dy \Big/ \int_{-\infty}^{\infty} f(X, y)\, dy, \qquad (28.2)$$

выражающим зависимость r-го момента величины y при фиксированном X от X. Аналогично, равенство

$$\nu_{rX} = \mathbf{M}[\{y - \mathbf{M}(y \mid X)\}^r \mid X] =$$

$$= \int_{-\infty}^{\infty} \{y - \mathbf{M}(y \mid X)\}^r f(X, y)\, dy \Big/ \int_{-\infty}^{\infty} f(X, y)\, dy \qquad (28.3)$$

выражает зависимость центрального момента величины y при фиксированном X от X.

Если в (28.3) $r = 2$, то это есть так называемая *скедастическая* (*scedastic*) кривая, выражающая зависимость от X дисперсии величины y при фиксированном X. Если нарисуем график, на котором каждому X будет соответствовать значение коэффициента асимметрии $\beta_{1X} = \mu^2_{3X}/\mu^3_{2X}$, то получим *клитическую* (*clitic*) кривую, а делая то же самое для величины $\beta_{2X} = \mu_{4X}/\mu^2_{2X}$, найдем *куртическую* (*kurtic*) кривую *). Приведенные определения не являются в действительности общепринятыми. Наиболее важное значение имеет регрессионная кривая, получающаяся при $r = 1$, т. е. кривая (28.1); и всякий раз, когда без уточнения будет упоминаться «регрессия», ее следует понимать как регрессию (28.1) для среднего значения.

В **26.5** мы видели, что иногда нам интересна регрессия x по y, так же как и регрессия y по x. Тогда мы имеем очевидные аналоги соотношений (28.2) и (28.3), в частности, соотношение, аналогичное (28.1), имеет вид

$$\mu'_{1Y} = \mathbf{M}(x \mid Y) = \int_{-\infty}^{\infty} x f(x, Y)\, dx \Big/ \int_{-\infty}^{\infty} f(x, Y)\, dx. \qquad (28.4)$$

28.3 В точности так же, как мы получали моменты из х. ф. без вычисления плотности распределения, можно находить регрессию любого порядка с помощью совместной х. ф. величин x

*) Возможно, более полезно было бы вместо моментов наносить на график относительно X семиинварианты величины y, хотя такая вещь, насколько нам известно, никогда не делалась.

и y без вычисления в явном виде совместной ф. п. $f(x, y)$. Запишем

$$f(x, y) = g(x) \cdot h_x(y), \tag{28.5}$$

где $g(x)$ есть маргинальное распределение величины x, а $h_x(y)$ — условное распределение величины y при фиксированном x*). Совместная х. ф. x и y равна

$$\varphi(t_1, t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(it_1x + it_2y)\, g(x)\, h_x(y)\, dx\, dy = \tag{28.6}$$

$$= \int_{-\infty}^{\infty} \exp(it_1x)\, g(x)\, \varphi_x(t_2)\, dx, \tag{28.7}$$

где

$$\varphi_x(t_2) = \int_{-\infty}^{\infty} \exp(it_2y)\, h_x(y)\, dy$$

есть условная х. ф. величины y при фиксированном x. Если, как в **28.2**, μ'_{rx} есть r-й момент величины y при фиксированном x, то

$$i^r \mu'_{rx} = \left[\frac{\partial^r}{\partial t_2^r} \varphi_x(t_2)\right]_{t_2=0}, \tag{28.8}$$

и, следовательно, из (28.7) и (28.8) получаем

$$\left[\frac{\partial^r}{\partial t_2^r} \varphi(t_1, t_2)\right]_{t_2=0} = i^r \int_{-\infty}^{\infty} \exp(it_1x)\, g(x)\, \mu'_{rx}\, dx. \tag{28.9}$$

Таким образом, согласно теореме обращения (**4.3**)

$$g(x)\, \mu'_{rx} = \frac{(-i)^r}{2\pi} \int_{-\infty}^{\infty} \exp(-it_1x) \left[\frac{\partial^r}{\partial t_2^r} \varphi(t_1, t_2)\right]_{t_2=0} dt_1. \tag{28.10}$$

Это и есть искомое выражение, с помощью которого можно получить регрессию любого порядка.

При $r = 1$ из (28.10) имеем

$$g(x)\, \mu'_{1x} = \frac{-i}{2\pi} \int_{-\infty}^{\infty} \exp(-it_1x) \left[\frac{\partial}{\partial t_2} \varphi(t_1, t_2)\right]_{t_2=0} dt_1. \tag{28.11}$$

*) Мы теперь не будем использовать X для обозначения фиксированного значения величины x.

28.4 Если существуют все семиинварианты, то мы имеем, по определению двумерных семиинвариантов (3.74),

$$\varphi(t_1, t_2) = \exp\left\{ \sum_{r,\, s=0}^{\infty} \varkappa_{rs} \frac{(it_1)^r}{r!} \frac{(it_2)^s}{s!} \right\},$$

где, по определению, $\varkappa_{00}$ равно нулю. Следовательно,

$$\left[\frac{\partial \varphi(t_1, t_2)}{\partial t_2}\right]_{t_2=0} = \left[i\varphi(t_1, t_2) \sum_{r=0}^{\infty} \sum_{s=1}^{\infty} \varkappa_{rs} \frac{(it_1)^r}{r!} \frac{(it_2)^{s-1}}{(s-1)!} \right]_{t_2=0} =$$

$$= i\varphi(t_1, 0) \sum_{r=0}^{\infty} \varkappa_{r1} \frac{(it_1)^r}{r!}, \quad (28.12)$$

Используя (28.12), перепишем (28.11) в виде

$$g(x)\, \mu'_{1x} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-it_1 x)\, \varphi(t_1, 0) \sum_{r=0}^{\infty} \varkappa_{r1} \frac{(it_1)^r}{r!}\, dt_1, \quad (28.13)$$

и если допустима перестановка операций интегрирования и суммирования, то (28.13) переходит в

$$g(x)\, \mu'_{1x} = \frac{1}{2\pi} \sum_{r=0}^{\infty} \frac{\varkappa_{r1}}{r!}\, i^r \int_{-\infty}^{\infty} t_1^r \exp(-it_1 x)\, \varphi(t_1, 0)\, dt_1. \quad (28.14)$$

Так как согласно теореме обращения

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-it_1 x)\, \varphi(t_1, 0)\, dt_1,$$

то, предполагая выполненными условия существования, получаем

$$(-D)^l g(x) \equiv (-1)^l \frac{d^l}{dx^l} g(x) = \frac{i^l}{2\pi} \int_{-\infty}^{\infty} t_1^l \exp(-it_1 x)\, \varphi(t_1, 0)\, dt_1. \quad (28.15)$$

Используя (28.15), находим из (28.14)

$$g(x)\, \mu'_{1x} = \sum_{r=0}^{\infty} \frac{\varkappa_{r1}}{r!} (-D)^r g(x). \quad (28.16)$$

Таким образом, для регрессии среднего значения величины y по x имеем

$$\mu'_{1x} = \sum_{r=0}^{\infty} \frac{\varkappa_{r1}}{r!} \frac{(-D)^r g(x)}{g(x)}. \quad (28.17)$$

Приведенный результат принадлежит Уикселлу (1934). Соотношение (28.17) справедливо, если существуют семиинварианты всех порядков и законна перестановка операций интегрирования и суммирования в (28.13). Это выполняется, в частности, тогда, когда $g(x)$ и все ее производные непрерывны внутри интервала изменения x и равны нулю на его концах.

Если $g(x)$ — нормированная нормальная плотность, то получаем частный случай соотношения (28.17):

$$\mu'_{1x} = \sum_{r=0}^{\infty} \frac{\varkappa_{r1}}{r!} H_r(x), \tag{28.18}$$

где $H_r(x)$ — полиномы Чебышева — Эрмита порядка r, определенные в (6.21).

Пример 28.1

Для двумерного нормального распределения

$$f(x, y) = (2\pi\sigma_1\sigma_2)^{-1}(1-\rho^2)^{-1/2}\exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right\}\right]$$

совместная х. ф. величин $\frac{x-\mu_1}{\sigma_1}$ и $\frac{y-\mu_1}{\sigma_2}$ (см. пример 15.1) равна

$$\varphi(t_1, t_2) = \exp\left\{-\frac{1}{2}(t_1^2 + t_2^2 + 2\rho t_1 t_2)\right\}.$$

Отсюда

$$\varkappa_{01} = 0, \quad \varkappa_{r1} = 0, \quad r > 1,$$

так что

$$\varkappa_{11} = \rho$$

есть единственный ненулевой семиинвариант в (28.17). Маргинальное распределение $g(x)$ является стандартным нормальным распределением, поэтому (28.17) совпадает с (28.18). Используя (6.23), находим

$$\mu'_{1x} = \varkappa_{11} H_1(x) = \rho x.$$

Это есть регрессия величины $(y-\mu_2)/\sigma_2$ по $(x-\mu_1)/\sigma_1$. Если теперь перейти к ненормированным величинам, то для регрессии y относительно x получим

$$\mathbf{M}(y \mid x) - \mu_2 = \frac{\rho\sigma_2}{\sigma_1}(x-\mu_1),$$

более общую форму первого из уравнений (16.46), где x и y были переставлены местами и $\mu_1 = \mu_2 = 0$.

Пример 28.2

Пусть дана выборка объема n из двумерной нормальной совокупности предыдущего примера. Рассмотрим совместное распределение величин

$$u = \frac{1}{2}\sum_{i=1}^{n}(x_i - \mu_1)^2/\sigma_1^2 \quad \text{и} \quad v = \frac{1}{2}\sum_{i=1}^{n}(y_i - \mu_2)^2/\sigma_2^2.$$

Из примера 26.1 легко найти, что совместная х. ф. величин u и v равна

$$\varphi(t_1, t_2) = \{(1-\theta_1)(1-\theta_2) - \rho^2\theta_1\theta_2\}^{-n/2}, \tag{28.19}$$

где $\theta_1 = it_1$, $\theta_2 = it_2$. Выразить в простой форме совместную ф. п. величин u и v невозможно, но регрессии можно определить и без нее. Из (28.19) находим

$$\left[\frac{\partial^r \varphi(t_1, t_2)}{\partial t_2^r}\right]_{t_2=0} = i^r\left[\frac{\partial^r \varphi}{\partial \theta_2^r}\right]_{\theta_2=0} = i^r\left(\frac{1}{2}n + r - 1\right)^{(r)}\frac{\{1-(1-\rho^2)\theta_1\}^r}{(1-\theta_1)^{(n/2)+r}}. \tag{28.20}$$

Таким образом, (28.10) и (28.20) дают

$$g(u)\,\mu'_{rv} =$$

$$= \left(\frac{1}{2}n + r - 1\right)^{(r)}\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(-\theta_1 u)\frac{\{\rho^2 + (1-\rho^2)(1-\theta_1)\}^r}{(1-\theta_1)^{(n/2)+r}}dt_1. \tag{28.21}$$

Из обращения х. ф., приведенного в примере 4.4, имеем

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\exp(-\theta_1 u)}{(1-\theta_1)^k}dt_1 = \frac{1}{\Gamma(k)}e^{-u}u^{k-1},$$

тогда как маргинальное распределение $g(u)$ (см. (11.8)) есть

$$g(u) = \frac{1}{\Gamma(n/2)}e^{-u}u^{(n/2)-1}.$$

Подставляя его в (28.21) и полагая поочередно $r = 1, 2$, находим

$$\mu'_{1v} = \frac{1}{2}n\left\{\rho^2\frac{v}{(n/2)} + (1-\rho^2)\right\} = \rho^2 v + \frac{1}{2}n(1-\rho^2) \tag{28.22}$$

и

$$\mu'_{2v} = \frac{1}{2}n\left(\frac{1}{2}n+1\right)\left\{\rho^4\frac{v^2}{(n/2)((n/2)+1)} + 2\rho^2(1-\rho^2)\frac{v}{(n/2)} + (1-\rho^2)^2\right\} =$$

$$= \rho^4 v^2 + 2\rho^2(1-\rho^2)v\left(\frac{1}{2}n+1\right) + (1-\rho^2)^2\frac{1}{2}n\left(\frac{1}{2}n+1\right),$$

так что

$$\mu_{2v} = \mu'_{2v} - (\mu'_{1v})^2 = (1-\rho^2)\left\{2\rho^2 v + \frac{1}{2} n (1-\rho^2)^2\right\}. \quad (28.23)$$

Соотношения (28.22) и (28.23) показывают, что регрессии среднего и дисперсии величины u по v линейны.

Критерии линейности регрессии

28.5 Пусть $\psi(t_1, t_2) = \log \varphi(t_1, t_2)$ — совместная п. ф. с. величин x и y. Сейчас мы докажем следующий факт: если регрессия величины y по x линейна, так что

$$\mu'_{1x} = \mathbf{M}(y \mid x) = \beta_0 + \beta_1 x, \quad (28.24)$$

то

$$\left[\frac{\partial \psi(t_1, t_2)}{\partial t_2}\right]_{t_2=0} = i\beta_0 + \beta_1 \frac{\partial \psi(t_1, 0)}{\partial t_1}; \quad (28.25)$$

и наоборот, если выполнено некоторое условие полноты, то (28.25) не только необходимо, но и достаточно для (28.24).

Используя (28.24), из (28.9) при $r = 1$ находим

$$\left[\frac{\partial \varphi(t_1, t_2)}{\partial t_2}\right]_{t_2=0} = i \int_{-\infty}^{\infty} \exp(it_1 x)\, g(x)(\beta_0 + \beta_1 x)\, dx = \quad (28.26)$$

$$= i\beta_0 \varphi(t_1, 0) + \beta_1 \frac{\partial}{\partial t_1} \varphi(t_1, 0). \quad (28.27)$$

Полагая в (28.27) $\psi = \log \varphi$ и деля обе части на $\varphi(t_1, 0)$, получаем (28.25).

Обратно, если имеет место соотношение (28.25), то, используя (28.9), перепишем его в виде

$$i \int_{-\infty}^{\infty} \exp(it_1 x)(\beta_0 + \beta_1 x - \mu'_{1x})\, g(x)\, dx = 0. \quad (28.28)$$

Теперь видим, что соотношение (28.28) влечет тождественно по x

$$\beta_0 + \beta_1 x - \mu'_{1x} = 0, \quad (28.29)$$

если только семейство $\exp(it_1 x) g(x)$ полно. Следовательно, мы получили (28.24).

28.6 Если все семиинварианты существуют, то, используя (28.12), находим из (28.25)

$$\sum_{r=0}^{\infty} \varkappa_{r1} \frac{(it_1)^r}{r!} = \beta_0 + \beta_1 \sum_{r=0}^{\infty} \varkappa_{r0} \frac{(it_1)^{r-1}}{(r-1)!}. \quad (28.30)$$

Приравнивая в (28.30) коэффициенты при t^r, имеем

$$(r=0) \qquad \varkappa_{01}=\beta_0+\beta_1\varkappa_{10} \tag{28.31}$$

(что очевидно из (28.24));

$$(r\geqslant 1) \qquad \varkappa_{r1}=\beta_1\varkappa_{r+1,0}. \tag{28.32}$$

Условие линейности регрессии (28.32) также принадлежит Уикселлу (1934). Соотношения (28.31) и (28.32), вместе взятые, необходимы и достаточны для (28.25), а отсюда (как и прежде, при условии полноты $g(x)$) и для условия линейности (28.24). Если, как в (28.27), мы выразим (28.25) через х. ф. φ, а не через ее логарифм ψ и проведем выкладки, ведущие к (28.32), то получим аналог соотношения (28.32) для центральных моментов:

$$\mu_{r1}=\beta_1\mu_{r+1,0}. \tag{28.33}$$

Если регрессия величины x по y тоже линейна, т. е. $x=\beta_0'+\beta_1'y$, то найдем также

$$\varkappa_{1r}=\beta_1'\varkappa_{0,r+1}, \qquad r\geqslant 1. \tag{28.34}$$

При $r=1$ соотношения (28.32) и (28.34) дают $\varkappa_{11}=\beta_1\varkappa_{20}=\beta_1'\varkappa_{02}$, откуда

$$\beta_1\beta_1'=\varkappa_{11}^2/(\varkappa_{20}\varkappa_{02})=\rho^2, \tag{28.35}$$

где ρ — коэффициент корреляции между x и y, что вновь совпадает с (26.17).

28.7 Наложим теперь дальнейшие ограничения на наши переменные. Предположим, что условное распределение величины y относительно своего среднего (которое, как и прежде, является функцией от фиксируемого значения x) одно и то же для любого x, т. е. только среднее значение величины y изменяется с изменением x. Мы будем в этом случае говорить, что y «имеет однородные ошибки». Таким образом, существует случайная величина ε такая, что

$$y=\mu_{1x}'+\varepsilon. \tag{28.36}$$

В частности, когда регрессия линейна, (28.36) представляется в форме

$$y=\beta_0+\beta_1x+\varepsilon. \tag{28.37}$$

Если y имеет однородные ошибки, то соотношение (28.5) переходит в

$$f(x, y)=g(x)\,h(\varepsilon), \tag{28.38}$$

где h есть теперь условное распределение величины ε. Обратно, (28.38) влечет однородные ошибки для y.

Не столь очевиден соответствующий результат для характеристических функций: если регрессия y по x линейна с одно-

родными ошибками, то совместная х. ф. величин x и y факторизуется в виде

$$\varphi(t_1, t_2) = \varphi_g(t_1 + t_2\beta_1)\,\varphi_h(t_2)\exp(it_2\beta_0), \tag{28.39}$$

где индексы у φ указывают на соответствующие ф. п. Для доказательства (28.39) заметим, что

$$\varphi(t_1, t_2) = \int\int \exp(it_1x + it_2y) f(x, y)\,dx\,dy =$$

$$= \int\int \exp\{it_1x + it_2(\beta_0 + \beta_1x + \varepsilon)\}\, g(x)\, h(\varepsilon)\,dx\,d\varepsilon =$$

$$= \int \exp\{i(t_1 + t_2\beta_1)x\}\, g(x)\,dx \int \exp(it_2\varepsilon)\, h(\varepsilon)\,d\varepsilon \exp(it_2\beta_0), \tag{28.40}$$

и (28.39) есть просто по-иному переписанное выражение (28.40). Отметим, что если $\beta_1 = 0$, то согласно (28.39) x и y *независимы*: линейность регрессии, однородные ошибки и нулевой коэффициент регрессии влекут независимость, как это интуитивно очевидно.

Характеризация двумерного нормального распределения

28.8 Теперь мы можем доказать замечательный результат: если регрессии y по x и x по y линейны с однородными ошибками, то x и y имеют двумерное нормальное распределение, за исключением случаев, когда (а) они независимы одна от другой или (б) они связаны функциональной зависимостью.

Предположим, что выполнены условия теоремы. Логарифмируя (28.39), получаем

$$\psi(t_1, t_2) = \psi_g(t_1 + t_2\beta_1) + \psi_h(t_2) + it_2\beta_0. \tag{28.41}$$

Аналогично, из регрессии величины x по y

$$\psi(t_1, t_2) = \psi_{g'}(t_2 + t_1\beta_1') + \psi_{h'}(t_1) + it_1\beta_0', \tag{28.42}$$

где штрихи использованы для отличения коэффициентов и распределений от таковых в (28.41). Приравняем (28.41) и (28.42) и рассмотрим последовательные степени t_1 и t_2, обозначая r-е семиинварианты распределений g, g', h, h' через $\varkappa_{r0}$, $\varkappa_{0r}$, λ_{r0}, λ_{0r} соответственно.

Первая степень:

$$\varkappa_{10}i(t_1 + t_2\beta_1) + \lambda_{10}it_2 + it_2\beta_0 = \varkappa_{01}i(t_2 + t_1\beta_1') + \lambda_{01}it_1 + it_1\beta_0'.$$

Приравнивая коэффициенты при t_1 и при t_2, получаем

$$\varkappa_{10} = \varkappa_{01}\beta_1' + \lambda_{01} + \beta_0', \tag{28.43}$$

$$\varkappa_{10}\beta_1 + \lambda_{10} + \beta_0 = \varkappa_{01}. \tag{28.44}$$

Мы всегда можем предполагать, что ошибки имеют нулевые средние, ибо в противном случае их средние можно было бы включить в β_0 и β_0'. Если, кроме того, измерять x и y от их средних, то из (28.43) и (28.44) находим

$$\beta_0' = \beta_0 = 0, \tag{28.45}$$

что очевидно из общих соображений.

Вторая степень:

$$\varkappa_{20}(t_1 + t_2\beta_1)^2 + \lambda_{20}t_2^2 = \varkappa_{02}(t_2 + t_1\beta_1')^2 + \lambda_{02}t_1^2.$$

Отсюда, приравнивая коэффициенты при t_1^2, t_1t_2 и t_2^2, получаем

$$\varkappa_{20} = \varkappa_{02}(\beta_1')^2 + \lambda_{02}, \tag{28.46}$$

$$\varkappa_{20}\beta_1 = \varkappa_{02}\beta_1', \tag{28.47}$$

$$\varkappa_{20}\beta_1^2 + \lambda_{20} = \varkappa_{02}. \tag{28.48}$$

Равенства (28.46)—(28.48) дают связь между g, h, g' и h'. В частности, из (28.47) вытекает, что отношение β_1/β_1' равно $\varkappa_{02}/\varkappa_{20}$, отношению исходных дисперсий.

Третья степень:

$$\varkappa_{30}\{i(t_1 + t_2\beta_1)\}^3 + \lambda_{30}(it_2)^3 = \varkappa_{03}\{i(t_2 + t_1\beta_1')\}^3 + \lambda_{03}(it_1)^3.$$

Коэффициенты при $t_1^2t_2$ и $t_1t_2^2$ дают нам

$$\varkappa_{30}\beta_1 = \varkappa_{03}(\beta_1')^2, \tag{28.49}$$

$$\varkappa_{30}\beta_1^2 = \varkappa_{03}\beta_1'. \tag{28.50}$$

Исключая временно возможности β_1, $\beta_1' = 0$ или $\beta_1\beta_1' = 1$, получаем из соотношений (28.49) и (28.50) равенства $\varkappa_{30} = \varkappa_{03} = 0$. Аналогично, если рассмотрим четвертые и более высокие степени, то обнаружим, что семиинварианты всех высших порядков $\varkappa_{r0}$, $\varkappa_{0r}$ должны равняться нулю. Тогда из уравнений такого типа, как получаются при сравнении членов с t_1^3, t_2^3, а именно

$$\varkappa_{30} = \varkappa_{03}(\beta_1')^3 + \lambda_{03}, \quad \varkappa_{30}\beta_1^3 + \lambda_{30} = \varkappa_{03},$$

следует, что для распределений h, h' семиинварианты порядка выше второго также обращаются в нуль. Таким образом, все распределения g, h, g', h' являются нормальными, а из (28.41) или (28.42) вытекает, что $\psi(t_1, t_2)$ является квадратичной формой по t_1, t_2, и, следовательно, величины x, y имеют двумерное нормальное распределение.

В тех исключительных случаях, которыми мы пренебрегли, наше рассуждение теряет силу. Если $\beta_1\beta_1' = 1$, то согласно

(28.35) корреляция между x и y равна ± 1, и x является строго линейной функцией от y (см. **26.9**). С другой стороны, когда β_1 или $\beta_1' = 0$, величины x, y независимы, как замечено в конце **28.7**. Это заканчивает доказательство теоремы *).

Многомерные обобщения

28.9 Теперь мы кратко укажем, как обобщить наши результаты на случай p регрессоров $x_1, x_2, \ldots, x_p$. Линейная регрессия тогда имеет вид

$$\mathbf{M}(y \mid x_1, \ldots, x_p) = \beta_0 + \beta_1 x + \ldots + \beta_p x_p. \qquad (28.51)$$

Записывая совместную ф. п. в форме, аналогичной (28.5),

$$f(y, x_1, \ldots, x_p) = g(\boldsymbol{x})\, h_{\boldsymbol{x}}(y),$$

где $g(\boldsymbol{x})$ есть p-мерное маргинальное распределение величин $x_1, \ldots, x_p$, находим, как и в (28.6), что

$$\varphi(u, t_1, \ldots, t_p) =$$

$$= \int \ldots \int \exp\left(iuy + i\sum_{j=1}^{p} t_j x_j\right) g(\boldsymbol{x})\, h_{\boldsymbol{x}}(y)\, d\boldsymbol{x}\, dy =$$

$$= \int \ldots \int \exp\left(i\sum_{j=1}^{p} t_j x_j\right) g(\boldsymbol{x})\, \varphi_{\boldsymbol{x}}(u)\, d\boldsymbol{x}. \qquad (28.52)$$

Последнее равенство аналогично (28.7). В точности, как в (28.8), $i^r \mu'_{r\boldsymbol{x}} = \left[\frac{\partial^r}{\partial u^r} \varphi_{\boldsymbol{x}}(u)\right]_{u=0}$, и, подобно (28.9), получаем соотношение

$$\left[\frac{\partial^r}{\partial u^r} \varphi(u, t_1, \ldots, t_p)\right]_{u=0} = i^r \int \exp\left(i \sum t_j x_j\right) g(\boldsymbol{x})\, \mu'_{r\boldsymbol{x}}\, d\boldsymbol{x}, \qquad (28.53)$$

позволяющее обобщить (28.10):

$$g(\boldsymbol{x})\, \mu'_{r\boldsymbol{x}} = \frac{(-i)^r}{2\pi} \int \exp\left(-i \sum t_j x_j\right) \left[\frac{\partial^r}{\partial u^r} \varphi(u, t_1, \ldots, t_p)\right]_{u=0} d\boldsymbol{t}. \qquad (28.54)$$

28.10 Читатель без труда сможет обобщить критерий линейности регрессии из **28.5**: если имеет место (28.51), то должно выполняться равенство

$$\left[\frac{\partial \psi(u, t_1, \ldots, t_p)}{\partial u}\right]_{u=0} = i\beta_0 + \sum_{j=1}^{p} \beta_j \frac{\partial}{\partial t_j} \psi(0, t_1, \ldots, t_p), \qquad (28.55)$$

*) Первый опубликованный результат такого рода принадлежит, по-видимому, Бернштейну (1928). Доказательство при более общих условиях можно найти у Ферона и Фуржо (1952).

обобщающее (28.25). Аналогично, обобщением критерия (28.32) служит равенство

$$\varkappa_{1,\, r_1,\, r_2,\, \ldots,\, r_p} = \beta_1 \varkappa_{0,\, r_1+1,\, r_2,\, \ldots,\, r_p} + \\ + \beta_2 \varkappa_{0,\, r_1,\, r_2+1,\, r_3,\, \ldots,\, r_p} + \ldots + \beta_p \varkappa_{0,\, r_1,\, \ldots,\, r_{p-1},\, r_p+1}. \quad (28.56)$$

Условие (28.38) однородности ошибок в этом случае имеет вид

$$f(y,\, x_1,\, \ldots,\, x_p) = g(\boldsymbol{x})\, h(t),$$

а (28.39), соответственно, превращается в

$$\varphi(u,\, t_1,\, \ldots,\, t_p) = \\ = \varphi_g(t_1 + u\beta_1,\, t_2 + u\beta_2,\, \ldots,\, t_p + u\beta_p)\, \varphi_h(u) \exp(iu\beta_0). \quad (28.57)$$

Наконец, обобщением **28.8** служит утверждение: если каждая из линейных регрессий в совокупности из p случайных величин имеет однородные ошибки, то эти величины имеют многомерное нормальное распределение, если только они не являются независимыми в совокупности или функционально связанными.

28.11 Можно получить аналогичные результаты, когда регрессия y по x полиномиальна, т. е. представима в форме

$$\mathbf{M}(y \mid x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \ldots + \beta_p x^p. \quad (28.58)$$

Однако ниже в этой главе мы увидим, что это лучше трактовать как частный случай p-регрессорной ситуации, в которой регрессоры функционально связаны. Поэтому любые требующиеся нам результаты о полиномиальной регрессии могут быть получены как частные случаи результатов из **28.9—10**. Например, условием, обеспечивающим (28.58), является равенство

$$\left[\frac{\partial \psi(u,\, t)}{\partial u}\right]_{u=0} = \\ = i\left\{\beta_0 + \beta_1 \frac{\partial \psi(0,\, t)}{\partial (it)} + \beta_2 \frac{\partial^2 \psi(0,\, t)}{\partial (it)^2} + \ldots + \beta_p \frac{\partial^p \psi(0,\, t)}{\partial (it)^p}\right\},$$

которое при $p = 1$ сводится к (28.25) (в несколько отличных обозначениях) и которое легко получить как частный случай из (28.55), заметив, что х. ф. величины x^r есть $\mathbf{M}\{\exp(itx^r)\}$, производная которой по t имеет вид

$$\mathbf{M}\{ix^r \exp(itx^r)\} = \frac{1}{i^{r-1}} \frac{\partial^r}{\partial t^r} \mathbf{M}\{\exp(itx)\}.$$

Общая модель линейной регрессии

28.12 Аналитическая теория регрессии, которую мы до сих пор обсуждали, представляет интерес для теоретической, но не для прикладной статистики, потому что она требует точного знания функциональной формы исходного распределения. Об-

ратимся теперь к общей модели линейной регрессии, которая часто используется на практике, поскольку она выражает собой упрощенные (но тем не менее в разумной мере реалистические) предположения. Это есть в действительности общая модель из **19.4**, в которой параметры θ служат коэффициентами регрессии. Таким образом, переписывая соотношение (19.8), получаем

$$y = X\beta + \varepsilon, \tag{28.59}$$

где β — вектор $(k \times 1)$ коэффициентов регрессии, X — матрица $(n \times k)$ известных коэффициентов (не случайных величин), а ε — вектор $(n \times 1)$ «ошибок», случайных величин (не обязательно нормальных) со средними и матрицей рассеяния

$$\mathbf{M}(\varepsilon) = 0, \quad V(\varepsilon) = \sigma^2 I. \tag{28.60}$$

Будем предполагать, что $n \geqslant k$ и $|X'X| \neq 0$.

Теперь применимы все результаты главы 19. Из (19.12) находим НК-оценку вектора β:

$$\hat{\beta} = (X'X)^{-1} X'y, \tag{28.61}$$

матрица рассеяния которой согласно (19.16) равна

$$V(\hat{\beta}) = \sigma^2 (X'X)^{-1}. \tag{28.62}$$

Из **19.6** вытекает, что она представляет собой несмещенную линейную МД-оценку вектора β. Наконец, из (19.41) находим, что s^2 есть несмещенная оценка для σ^2, причем

$$(n-k)\, s^2 = (y - X\hat{\beta})'(y - X\hat{\beta}) \equiv y'y - \hat{\beta}'X'y. \tag{28.63}$$

Величина s^2 является суммой квадратов остатков, деленной на число наблюдений без числа оцениваемых параметров.

Эту модель мы уже применяли к регрессионным ситуациям в **26.8** и **27.15**.

Смысл слова «линейный»

28.13 Прежде чем переходить к дальнейшему, полезно подчеркнуть смысл прилагательного «линейный» в общей модели линейной регрессии (28.59): она *линейна относительно параметров* β_i, но не обязана быть линейной относительно иксов. Действительно, как мы уже отмечали, элементами матрицы X может быть любое множество известных констант, связанных одна с другой произвольным образом. С другой стороны, начиная с **28.11**, слова «линейная регрессия» означали, что условное математическое ожидание величины y является линейной функцией регрессоров $x_1, \ldots, x_p$. С точки зрения нашего настоящего

(по методу наименьших квадратов) анализа последнее (возможно, более «естественное») определение линейности не относится к делу; существенна только линейность относительно параметров. Таким образом, модель линейной регрессии включает в себя все виды «полиномиальной» или «криволинейной» зависимости y от $x_1, \ldots, x_p$. Например, непосредственно полиномиальная зависимость

$$y_j = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{1j}^2 + \ \ldots\ + \beta_k x_{1j}^k + \varepsilon_j, \qquad j = 1, 2, \ldots, n, \tag{28.64}$$

линейна относительно β и, следовательно, является специальным случаем соотношения (28.59) (см. замечания в **28.11**). Аналогично, в случае нескольких переменных соотношение

$$y_j = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{1j}^2 + \beta_3 x_{2j} + \beta_4 x_{2j}^2 + \beta_5 x_{1j} x_{2j} + \varepsilon_j, \quad j = 1, 2, \ldots, n, \tag{28.65}$$

представляет собой модель линейной регрессии. С другой стороны, соотношение

$$y_j = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \beta_1^2 x_{3j} + \varepsilon_j, \qquad j = 1, 2, \ldots, n,$$

не является таковым, поскольку в него входят и β_1 и β_1^2.

В линейную модель могут входить не только полиномиальные функции. Так, модель

$$y_j = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{1j}^2 + \beta_3 \sin x_{1j} \cos x_{2j} + \varepsilon_j$$

является линейной.

28.14 При таком понимании линейного регрессионного анализа мы видим, что он сводится к математической задаче обращения матрицы, состоящей из сумм квадратов и произведений регрессоров, $\boldsymbol{X'X}$. Обращение требуется для оценки (28.61) и для вычисления матрицы рассеяния оценок с помощью (28.62) и (28.63). При этом не возникает никакой новой статистической задачи.

Кокрэн (1938) и Кейб (1963) приводят формулы, дающие поправки на случай, когда одна или две из первоначальных величин x отсутствуют либо одна или две новых добавлено. Хадсон (1966) обсуждал общую задачу подгонки с помощью метода НК кусочно-линейных кривых, точки соединения которых должны быть оценены.

Ортогональный регрессионный анализ

28.15 Очевидно, удобно применять регрессионный анализ, когда оценки $\hat{\beta}_i$ некоррелированы. Действительно, если ε_j — нормальные величины, то таковыми будут и $\hat{\beta}_i$, так как они являются линейными функциями от $\boldsymbol{y}$, и тогда отсутствие корреляции влечет независимость. Регрессионный анализ с некорре-

лированными оценками называется *ортогональным*. Поскольку теперь регрессоры — не случайные величины, а константы, которые выбираются в экспериментальной работе, то мы можем поставить вопрос нового типа: можно ли и если можно, то как выбрать элементы матрицы $\boldsymbol{X}$, чтобы оценки $\hat{\beta}_i$ были некоррелированы?

Этот вопрос возникает в теории планирования эксперимента, а детальное обсуждение задач планирования мы отложим до третьего тома. Однако из (28.62) видно, что анализ ортогонален тогда и только тогда, когда матрица $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ диагональна. А $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ диагональна, лишь когда диагональна $\boldsymbol{X}'\boldsymbol{X}$. Таким образом, для получения ортогонального анализа мы должны выбрать элементы матрицы $\boldsymbol{X}$ такими, чтобы $\boldsymbol{X}'\boldsymbol{X}$ была диагональной матрицей. Отсюда немедленно вытекает условие

$$\sum_{j=1}^{n} x_{ij}x_{hj} = 0, \qquad i \neq h. \tag{28.66}$$

Диагональные элементы $\boldsymbol{X}'\boldsymbol{X}$ равны

$$(\boldsymbol{X}'\boldsymbol{X})_{ii} = \sum_{j=1}^{n} x_{ij}^2;$$

следовательно, соответствующие обратные элементы равны

$$[(\boldsymbol{X}'\boldsymbol{X})^{-1}]_{ii} = 1 \Big/ \sum_{j=1}^{n} x_{ij}^2. \tag{28.67}$$

В этом случае соотношения (28.61) и (28.62) выглядят особенно просто.

Полиномиальная регрессия: ортогональные полиномы

28.16 При полиномиальной зависимости y от x, как в (28.64), матрица $\boldsymbol{X}'\boldsymbol{X}$ не может быть диагональной, поскольку недиагональные элементы являются суммами степеней одной и той же величины x. Однако можно выбрать полиномы степени i относительно x, скажем $\varphi_i(x)$ $(i = 0, 1, 2, \ldots, k)$, которые будут взаимно ортогональны. Тогда (28.64) заменится соотношением

$$y_j = \alpha_0\varphi_0(x_j) + \alpha_1\varphi_1(x_j) + \ldots + \alpha_k\varphi_k(x_j) + \varepsilon_j, \qquad j = 1, 2, \ldots, n, \tag{28.68}$$

которое можно переписать в матричной форме как $\boldsymbol{y} = \boldsymbol{\Phi\alpha} + \boldsymbol{\varepsilon}$.

Величины α образуют новое множество параметров (зависящих от исходных β), в терминах которых значительно удобнее

работать. Действительно, теперь из (28.67) следует равенство

$$[(\mathbf{\Phi}'\mathbf{\Phi})^{-1}]_{ii} = 1 \Big/ \sum_{j=1}^{n} \varphi_i^2(x_j), \tag{28.69}$$

которое можно использовать в (28.62). Далее, учитывая (28.69), получаем из (28.61)

$$\hat{\alpha}_i = [(\mathbf{\Phi}'\mathbf{\Phi})^{-1}]_{ii} (\mathbf{\Phi}'\boldsymbol{y})_i = \sum_j y_i \varphi_i(x_j) \Big/ \sum_j \varphi_i^2(x_j). \tag{28.70}$$

Таким образом, каждая оценка $\hat{\alpha}_i$ зависит только от соответствующего полинома. Это очень удобно при «подгонке» полиномиальной регрессии, степень которой заранее не определена: мы шаг за шагом увеличиваем значение k до тех пор, пока не получим достаточно хорошее согласие. Если бы мы использовали неортогональную модель регрессии (28.64), то при добавлении нового члена x^k, увеличивающего степень полинома, нужно было бы пересчитывать все множество оценок $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_{k-1}$.

Конечно, если переписать оцененную регрессию $y = \sum_{i=0}^{k} \hat{\alpha}_i \varphi_i(x)$ в виде $y = \sum_{i=0}^{k} \hat{\beta}_i x^i$, то $\hat{\beta}_i$ будут в точности теми оценками, которые мы должны были бы получить, хотя и менее удобным способом, непосредственно из (28.64). Это вытекает из того факта, что оба метода минимизируют одну и ту же сумму квадратов остатков.

Используя (28.70), получаем в ортогональном случае из (28.63)

$$(n-k)\, s^2 = \boldsymbol{y}'\boldsymbol{y} - \hat{\boldsymbol{\alpha}}'\mathbf{\Phi}'\boldsymbol{y} =$$

$$= \sum_{j=1}^{n} y_j^2 - \sum_{i=1}^{k} \left[\left\{ \sum_{j=1}^{n} y_j \varphi_i(x_j) \right\}^2 \Big/ \sum_{j=1}^{n} \varphi_i^2(x_j) \right] = \tag{28.71}$$

$$= \sum_{j=1}^{n} y_j^2 - \sum_{i=1}^{k} \hat{\alpha}_i^2 \sum_{j=1}^{n} \varphi_i^2(x_j). \tag{28.72}$$

Эти очень простые выражения для суммы квадратов остатков от подгоняемой регрессии позволяют быстро вычислить дополнительное уменьшение остаточной дисперсии, вызываемое увеличением степени k подгоняемого полинома.

28.17 Рассмотрим теперь, как находить ортогональные полиномы $\varphi_i(x)$. Нам требуется, чтобы

$$\sum_{j=1}^{n} \varphi_i(x_j)\, \varphi_h(x_j) = 0, \qquad i, h = 0, 1, 2, \ldots, k; \qquad i \neq h, \tag{28.73}$$

где

$$\varphi_l(x) = \sum_{r=0}^{l} c_{lr} x^r. \qquad (28.74)$$

В (28.74) входит $l+1$ коэффициентов c_{lr}, а следовательно, все вместе полиномы φ_l содержат $\sum_{l=0}^{k}(l+1) = \frac{1}{2}(k+1)(k+2)$ коэффициентов, подлежащих определению. Для этой цели соотношения (28.73) дают только $\frac{1}{2}k(k+1)$ уравнений. Для определения оставшихся $k+1$ постоянных потребуем, чтобы $c_{ii} = 1$ при всех i. Тогда из (28.74) сразу получаем

$$\varphi_0(x) = c_{00} \equiv 1 \qquad (28.75)$$

тождественно по x. Теперь соотношения (28.73) определяют для каждого $\varphi_i(x)$, $i > 0$, коэффициенты c_{ir} с точностью до постоянного множителя, скажем λ_{in}. При $h = k$ из (28.73) и (28.74) получаем уравнения

$$\sum_{j=1}^{n} \sum_{r=0}^{l} c_{ir} x_j^r \sum_{s=0}^{k} c_{ks} x_j^s = 0, \qquad i \neq k,$$

или

$$\sum_{r=0}^{l} c_{ir} \sum_{s=0}^{k} c_{ks} \mu'_{r+s} = 0, \qquad i \neq k, \qquad (28.76)$$

где μ'_p есть p-й момент множества иксов. Так как (28.76) выполняется для всех $i = 0, 1, 2, \ldots, k-1$, то мы должны иметь равенства

$$\sum_{s=0}^{k} c_{ks} \mu'_{r+s} = 0, \qquad r = 0, 1, \ldots, k-1. \qquad (28.77)$$

Выпишем определитель

$$|M_k| = \begin{vmatrix} \mu'_0 & \mu'_1 & \cdots & \mu'_{k-1} & \mu'_k \\ \mu'_1 & \mu'_2 & \cdots & \mu'_k & \mu'_{k+1} \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \mu'_{k-1} & \mu'_k & \cdots & \mu'_{2k-2} & \mu'_{2k-1} \\ \mu'_k & \mu'_{k+1} & \cdots & \mu'_{2k-1} & \mu'_{2k} \end{vmatrix},$$

и пусть $|M_k^{u,v}|$ — минор элемента, стоящего на пересечении u-й строки и v-го столбца в определителе $|M_k|$. Тогда решение

системы (28.77) (с учетом предположения $c_{kk} = 1$) есть

$$c_{ks} = |M_k^{k+1,\,s+1}| / |M_{k-1}|, \quad s = 0, 1, 2, \ldots, k-1. \tag{28.78}$$

Таким образом, из (28.74) и (28.78) следует, что

$$\varphi_k(x) = \frac{1}{|M_{k-1}|} \begin{vmatrix} \mu_0' & \mu_1' & \ldots & \mu_k' \\ \mu_1' & \mu_2' & \ldots & \mu_{k+1}' \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \mu_{k-1}' & \mu_k' & \ldots & \mu_{2k-1}' \\ 1 & x & \ldots & x^k \end{vmatrix}. \tag{28.79}$$

Соотношение (28.79) применимо для нахождения полинома, каково бы ни было k. Разумеется, $\mu_0' \equiv 1$. Если измерять иксы от их средних значений, так, чтобы $\mu_1' = 0$, то штрихи можно опустить. Следует заметить, что если в (28.79) у определителя, стоящего в числителе, вычеркнуть последние строку и столбец, то он совпадет со знаменателем.

К примеру, находим

$$\varphi_1(x) = \frac{\begin{vmatrix} 1 & 0 \\ 1 & x \end{vmatrix}}{|1|} = x, \tag{28.80}$$

$$\varphi_2(x) = \frac{\begin{vmatrix} 1 & 0 & \mu_2 \\ 0 & \mu_2 & \mu_3 \\ 1 & x & x^2 \end{vmatrix}}{\begin{vmatrix} 1 & 0 \\ 0 & \mu_2 \end{vmatrix}} = x^2 - \frac{\mu_3}{\mu_2} x - \mu_2 \tag{28.81}$$

и т. д. В упражнении 28.23 дан более простой, рекуррентный способ нахождения полиномов.

Случай равноотстоящих значений иксов

28.18 Ортогональные полиномы находят наиболее важное применение в тех ситуациях, когда регрессорная переменная x принимает значения через равные промежутки. Такая ситуация возникает, когда наблюдения производятся в последовательные моменты времени или группированы в классы, имеющие одинаковую ширину. Если есть n таких равноотстоящих значений x, то, измеряя от их среднего значения и принимая интервал между ними за единицу, получим в качестве рабочих значений x: $-\frac{1}{2}(n-1)$, $-\frac{1}{2}(n-3)$, $-\frac{1}{2}(n-5)$, $\ldots$, $\frac{1}{2}(n-3)$, $\frac{1}{2}(n-1)$.

Для этого простого случая значения моментов (28.79) могут быть вычислены в явном виде. Действительно, за исключением среднего, которое принято за начало координат, они являются моментами совокупности первых n натуральных чисел и могут быть определены по семиинвариантам, приведенным в упражнении 3.23. Из симметрии нечетные моменты равны нулю; четные моменты равны

$$\mu_2 = (n^2 - 1)/12,$$
$$\mu_4 = \mu_2 (3n^2 - 7)/20,$$
$$\mu_6 = \mu_2 (3n^4 - 18n^2 + 31)/112$$

и т. д. Подставляя эти и моменты более высокого порядка в (28.79), получаем для первых шести полиномов выражения

$$\left.\begin{aligned}
&\varphi_0(x) \equiv 1,\\
&\varphi_1(x) = \lambda_{1n} x,\\
&\varphi_2(x) = \lambda_{2n}\left\{ x^2 - \frac{1}{12}(n^2 - 1)\right\},\\
&\varphi_3(x) = \lambda_{3n}\left\{ x^3 - \frac{1}{20}(3n^2 - 7)\, x\right\},\\
&\varphi_4(x) = \lambda_{4n}\left\{ x^4 - \frac{1}{14}(3n^2 - 13)\, x^2 + \frac{3}{560}(n^2-1)(n^2-9)\right\},\\
&\varphi_5(x) = \lambda_{5n}\left\{ x^5 - \frac{5}{18}(n^2-7)\, x^3 + \frac{1}{1008}(15n^4 - 230n^2 + 407)\, x\right\},\\
&\varphi_6(x) = \lambda_{6n}\left\{ x^6 - \frac{5}{44}(3n^2 - 31)\, x^4 + \right.\\
&\left. + \frac{1}{176}(5n^4 - 110n^2 + 329)\, x^2 - \frac{5}{14\,874}(n^2-1)(n^2-9)(n^2-25)\right\}.
\end{aligned}\right\} \qquad (28.82)$$

Аллан (1930) производит также $\varphi_i(x)$ для $i = 7, 8, 9, 10$. Следуя Фишеру (1921b), произвольные константы λ_{in} в (28.82), о которых говорилось после (28.75), определяют (из соображений удобства) так, чтобы значения $\varphi_i(x_j)$ были целыми при всех $j = 1, 2, \ldots, n$. Можно заметить, что

$$\varphi_{2i}(x) = \varphi_{2i}(-x) \quad \text{и} \quad \varphi_{2i-1}(x) = -\varphi_{2i-1}(-x),$$

т. е. полиномы четных степеней являются четными функциями, а полиномы нечетных степеней — нечетными функциями.

Таблицы ортогональных полиномов

28.19. *Biometrika Tables* содержит $\varphi_i(x_j)$ для всех j, $n = 3\,(1)\,52$ и $i = 1\,(1)\min(6,\ n-1)$, а также значения λ_{in} и $\sum_{j=1}^{n} \varphi_i^2(x_j)$ *).

*) Эти таблицы приведены в книге *Таблицы математической статистики.* (*Прим. ред.*)

В *Таблицах* Фишера и Иэйтса приведены $\varphi_i(x_j)$ (у них ξ_i'), λ_{in} и $\left[\sum_{j=1}^{n}\varphi_i^2(x_j)\right]$ для всех j, $n=3\,(1)\,75$ и $i=1\,(1)\,\min(5,\,n-1)$.

В *Biometrika Tables* можно найти ссылки на более обширные таблицы, доходящие до $i=9$, $n=52$ (Ван дер Рейден) и до $i=5$, $n=104$ (Андерсон и Хаусмэн).

28.20 Об ортогональных полиномах существует обширная литература. За теоретическими деталями читатель может обратиться к работе Фишера (1921 b), который первым применил их к полиномиальной регрессии, к работе Аллана (1930), а также к трем работам Эйткина (1933). Назовем несколько более современных источников. У Раштона (1951) обсуждается случай неодинаково отстоящих значений x, Кокс (1958) приводит краткий вывод общих ортогональных полиномов с использованием определителей. Задачи с группированными данными рассматривались Гестом (1954, 1956).

Приведем теперь один пример подгонки ортогональных полиномов в равноотстоящем случае.

Пример 28.3

Первые два столбца таблицы 28.1 показывают численность населения Англии и Уэлса по десятилетним переписям с 1811 по

Таблица 28.1

Год	Население (в миллионах) y	$n=13$: $\frac{\text{Год}-1871}{10}=x=\varphi_1(x)$	$\varphi_2(x)$	$\varphi_3(x)$	$\varphi_4(x)$
1811	10,16	−6	22	−11	99
1821	12,00	−5	11	0	−66
1831	13,90	−4	2	6	−96
1841	15,91	−3	−5	8	−54
1851	17,93	−2	−10	7	11
1861	20,07	−1	−13	4	64
1871	22,71	0	−14	0	84
1881	25,97	1	−13	−4	64
1891	29,00	2	−10	−7	11
1901	32,53	3	−5	−8	−54
1911	36,07	4	2	−6	−96
1921	37,89	5	11	0	−66
1931	39,95	6	22	11	99
$\sum y=314{,}09$	$\lambda_{1,13}$:	1	1	1/6	7/12
	$\sum_{j=1}^{13}\varphi_i^2(x_j)$:	182	2002	572	68 068

1931 г. Ясно, что эти наблюдения коррелированы, и поэтому регрессионная модель (28.64) не является строго соответствующей. Но с чисто иллюстративными целями мы проведем процесс подгонки. Здесь $n = 13$. Используя таблицу 47 *Biometrika Tables*, находим значения, стоящие в последних четырех столбцах таблицы 28.1. Имеем

$$\sum y_j \Phi_0(x_j) = \sum y_j = 314{,}09,$$

$$\sum y_j \Phi_1(x_j) = 474{,}77,$$

$$\sum y_j \Phi_2(x_j) = 123{,}19,$$

$$\sum y_j \Phi_3(x_j) = -39{,}38,$$

$$\sum y_j \Phi_4(x_j) = -374{,}30.$$

Следовательно, из (28.70) получаем

$$\hat{\alpha}_0 = 314{,}09/13 = 24{,}1608,$$
$$\hat{\alpha}_1 = 474{,}77/182 = 2{,}60863,$$
$$\hat{\alpha}_2 = 123{,}19/2002 = 0{,}0615335,$$
$$\hat{\alpha}_3 = -39{,}38/572 = -0{,}0688462,$$
$$\hat{\alpha}_4 = -374{,}30/68\,068 = -0{,}00549891.$$

Тогда согласно (28.68) и (28.82) ортогональная полиномиальная регрессия четвертой степени y по x имеет вид

$$y = 24{,}1608 + 2{,}60863x + 0{,}0615335(x^2 - 14) -$$
$$- 0{,}0688462\left\{\frac{1}{6}(x^3 - 25x)\right\} - 0{,}00549891\left\{\frac{7}{12}\left(x^4 - \frac{247}{7}x^2 + 144\right)\right\}.$$

Если бы мы, собирая подобные члены, привели правую часть к виду

$$y = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2 + \hat{\beta}_3 x^3 + \hat{\beta}_4 x^4,$$

то коэффициенты $\hat{\beta}_i$ были бы в точности те же, как если бы мы использовали (28.64) вместо ортогональной формы (28.68). Преимущество последней, кроме ее вычислительной простоты, состоит в том, что можно легко изучать улучшение в «подгонке» уравнения регрессии с ростом его степени. Для этой цели требуется лишь вычислить

$$\sum_j y_j^2 = 8839{,}939$$

и найденные ранее величины подставить в (28.72). Таким образом, имеем:

Общая сумма квадратов		8839,939
Уменьшение, обусловленное $\hat{a}_0$, равно $\hat{a}_0^2\sum\varphi_0^2=(24{,}1608)^2\cdot 13$	$=$	7588,656
	Остаток:	1251,283
» » $\hat{a}_1$ » $\hat{a}_1^2\sum\varphi_1^2=(2{,}60863)^2\cdot 182$	$=$	1238,497
	Остаток:	12,786
» « $\hat{a}_2$ » $\hat{a}_2^2\sum\varphi_2^2=(0{,}0615335)^2\cdot 2002=$		7,580
	Остаток:	5,206
» « $\hat{a}_3$ » $\hat{a}_3^2\sum\varphi_3^2=(0{,}0688462)^2\cdot 572=$		2,711
	Остаток:	2,495
» » $\hat{a}_4$ » $\hat{a}_4^2\sum\varphi_4^2=(0{,}00549891)^2\cdot 68068=$		2,058
	Остаток:	0,437

Очевидно, что линии третьего и четвертого порядка «подогнаны» хорошо. Они изображены на рис. 28.1.

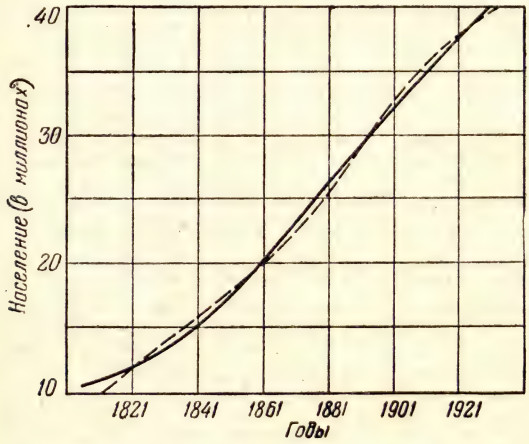


Рис. 28.1. Полиномы третьей (сплошная линия) и четвертой (пунктирная линия) степеней, подогнанные по данным таблицы 28.1.

Нет нужды предостерегать читателя против опасности экстраполяции на основе подогнанной регрессии, даже близкой, если она не имеет теоретической основы. В данном случае, например, можно наглядно убедиться в том, что значение, «предсказанное» регрессией четвертой степени для 1951 г. ($x=8$), значительно меньше того, что в действительности дала перепись населения в этом году (43,7 миллиона).

Доверительные интервалы и критерии для параметров линейной модели

28.21 В **28.12** мы обсуждали точечное оценивание параметров β, σ^2 общей линейной модели регрессии (28.59). Если теперь предположить, что вектор ошибок ε распределен *нормально* (это предположение сохранится до конца данной главы), то можно строить доверительные интервалы или, соответственно, проверять гипотезы для любой компоненты параметрического вектора β. Все они будут линейными гипотезами в смысле главы 24, а все критерии будут критериями ОП.

Любая оценка $\hat{\beta}_i$ есть линейная функция от y и, следовательно, распределена нормально со средним β_i и дисперсией, равной согласно (28.62)

$$\mathsf{D}(\hat{\beta}_i) = \sigma^2 [(X'X)^{-1}]_{ii}. \tag{28.83}$$

(Если анализ ортогональный, то в (28.83) используется (28.67).) Из **19.11** вытекает, что s^2, оценка для σ^2, определенная в (28.63), распределена независимо от $\hat{\beta}$ (и, следовательно, от любой компоненты $\hat{\beta}$), а величина $(n-k)s^2/\sigma^2$ имеет χ^2-распределение с $\nu = n-k$ степенями свободы. Отсюда сразу следует, что статистика

$$t = (\hat{\beta}_i - \beta_i) / \{s^2 [(X'X)^{-1}]_{ii}\}^{1/2}, \tag{28.84}$$

являясь отношением нормированной нормальной величины к корню квадратному из независимой от нее величины χ^2/ν, подчинена t-распределению Стьюдента с $\nu = n-k$ степенями свободы. Это позволяет нам строить доверительные интервалы для параметра β_i или проверять гипотезы, касающиеся его значений. Центральный доверительный интервал с коэффициентом доверия $(1-\alpha)$ имеет вид

$$\hat{\beta}_i \pm t_{1-(\alpha/2)} \{s^2 [(X'X)^{-1}]_{ii}\}^{1/2}, \tag{28.85}$$

где $t_{1-(\alpha/2)}$ — значение t-величины Стьюдента с ν степенями свободы, для которого ее функция распределения

$$F(t_{1-(\alpha/2)}) = 1 - \frac{1}{2}\alpha.$$

Поскольку здесь мы проверяем линейную гипотезу, то критерий, основанный на (28.84), является специальным случаем приведенного в **24.28** общего F-критерия дисперсионного отношения для линейной гипотезы: у нас сейчас только одно ограничение, и F-критерий сводится к t^2-критерию, соответствующему центральному доверительному интервалу (28.85).

Доверительные интервалы для условного математического ожидания y

28.22 Предположим, что, подогнав модель линейной регрессии к n наблюдениям, мы хотим оценить условное математическое ожидание y при фиксированных значениях каждого из k регрессоров $x_1, \ldots, x_k$. Если записать эти значения в виде $(1 \times k)$-вектора $\boldsymbol{x}^0$, то из **19.6** сразу вытекает, что при фиксированном $\boldsymbol{x}^0$ несмещенной оценкой с минимальной дисперсией для математического ожидания y является

$$\hat{y} = (\boldsymbol{x}^0)' \hat{\boldsymbol{\beta}} \tag{28.86}$$

и что согласно **19.6** и (28.62) ее дисперсия равна

$$\mathsf{D}\hat{y} = (\boldsymbol{x}^0)' \boldsymbol{V}(\hat{\boldsymbol{\beta}}) \boldsymbol{x}^0 = \sigma^2 (\boldsymbol{x}^0)' (\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{x}^0. \tag{28.87}$$

Так же как в 28.21, мы оценим выборочную дисперсию (28.87), заменяя σ^2 на s^2, и получим доверительные границы, используя t-распределение Стьюдента, которому в данном случае подчиняется статистика

$$t = \{\hat{y} - \mathsf{M}(y \mid \boldsymbol{x}^0)\} / \left\{s^2 (\boldsymbol{x}^0)' (\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{x}^0\right\}^{1/2}, \tag{28.88}$$

где, как и прежде, $\nu = n - k$.

Доверительные интервалы для последующего значения величины y: интервалы предсказания

28.23 Предположим, что регрессионная модель подогнана. Тогда результаты из **28.22** могут быть применены для получения доверительного интервала, содержащего последующее $(n+1)$-е значение y_{n+1} величины y. Если $\boldsymbol{x}^0$ — заданные значения регрессоров, при которых должно наблюдаться y_{n+1}, то соотношение (28.86) дает нам, как и прежде, несмещенную оценку математического ожидания y_{n+1}:

$$\hat{y} = (\boldsymbol{x}^0)' \hat{\boldsymbol{\beta}}. \tag{28.89}$$

Поскольку y_{n+1} имеет дисперсию σ^2, то дисперсия $y_{n+1} - \hat{y}$ увеличивается на эту величину по сравнению с (28.87), т. е.

$$\mathsf{D}(y_{n+1} - \hat{y}) = \sigma^2 \left\{(\boldsymbol{x}^0)' (\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{x}^0 + 1\right\}. \tag{28.90}$$

Подставляя s^2 вместо σ^2, находим оценку выборочной дисперсии и таким образом получаем статистику Стьюдента

$$t = \{y_{n+1} - \hat{y}\} / \left[s^2 \left\{(\boldsymbol{x}^0)' (\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{x}^0 + 1\right\}\right]^{1/2} \tag{28.91}$$

с $\nu = n - k$, на основе которой строятся искомые доверительные интервалы.

Аналогично, если должно быть сделано N последующих наблюдений над y при том же самом $\boldsymbol{x}^0$, то соотношения (28.89)—(28.91) имеют место для оценки среднего значения $\bar{y}_N$, которое должно наблюдаться, с очевидной заменой единицы в фигурных скобках (28.90) и (28.91) на $1/N$, так как дополнительная дисперсия равна теперь σ^2/N.

Доверительные интервалы для последующих значений, рассмотренные в этом пункте, иногда называются *интервалами предсказания*. Но мы должны помнить, что эти «предсказания» опираются на предположение, что линейная модель, подогнанная по n предшествующим наблюдениям, справедлива также и при последующих наблюдениях, т. е. что в модели не происходит структурных изменений.

Пример 28.4

Для простого случая

$$y_j = \beta_1 + \beta_2 x_j + \varepsilon_j, \qquad j = 1, 2, \ldots, n, \tag{28.92}$$

в примерах 19.3 и 19.6 мы видели, что

$$\hat{\beta}_2 = \sum_j (y_j - \bar{y})(x_j - \bar{x}) \Big/ \sum_j (x_j - \bar{x})^2, \qquad \hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x},$$

$$s^2 = \frac{1}{n-2} \sum_j \{y_j - (\hat{\beta}_1 + \hat{\beta}_2 x_j)\}^2$$

и

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \frac{1}{\sum_j (x_j - \bar{x})^2} \begin{pmatrix} \sum x^2/n & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix}.$$

Здесь $\boldsymbol{x}^0$ есть двухкомпонентный вектор $\begin{pmatrix} 1 \\ x_0 \end{pmatrix}$, и мы можем строить доверительные интервалы для β_1, β_2, $\mathbf{M}(y|\boldsymbol{x}^0)$ и y_{n+1}, используя соотношения (28.84), (28.88) и (28.91). В каждом случае мы имеем статистику Стьюдента с $n-2$ степенями свободы.

(а) Заметим, что анализ является ортогональным тогда и только тогда, когда $\bar{x} = 0$; поэтому для получения ортогональности здесь достаточно изменить начало отсчета переменной x. Аналогично, дисперсии оценок (диагональные элементы их дисперсионной матрицы) минимальны, если $\bar{x} = 0$, а $\sum x^2$ принимает максимально возможное значение. Ортогональность и минимальность выборочных дисперсий будут, следовательно, обеспечены, если выбрать x_j так, что (предполагая n четным)

$$x_1, x_2, \ldots, x_{n/2} = +a, \quad x_{(n/2)+1}, x_{(n/2)+2}, \ldots, x_n = -a,$$

где константа a — наибольшая из возможных. Это соответствует следующему интуитивно очевидному факту: если мы уверены

в том, что зависимость y от x линейна с постоянной дисперсией, то эффективнее всего «закреплять» прямую в концевых точках. Однако если бы зависимость была не линейной, то, производя все наблюдения только при двух значениях x, мы не могли бы этого обнаружить, и поэтому полезно распределить значения x более равномерно по интервалу. Всегда хорошо иметь возможность в процессе анализа проверять структурные предположения нашей модели.

(б) Доверительный интервал для $\mathbf{M}(y|x^0)$ согласно (28.88) есть

$$(x^0)'\hat{\beta} \pm t_{1-\alpha/2}\left\{\frac{s^2}{\sum(x-\bar{x})^2}\begin{pmatrix}1\\x^0\end{pmatrix}'\begin{pmatrix}\sum x^2/n & -\bar{x}\\ -\bar{x} & 1\end{pmatrix}\begin{pmatrix}1\\x^0\end{pmatrix}\right\}^{1/2} =$$

$$= (\hat{\beta}_1 + \hat{\beta}_2 x^0) \pm t_{1-\alpha/2}\left\{s^2\left(\frac{1}{n} + \frac{(x^0-\bar{x})^2}{\sum(x-\bar{x})^2}\right)\right\}^{1/2}. \qquad (28.93)$$

Если рассматривать это как функцию от x^0, то видно, что (28.93) определяет две ветви гиперболы, диаметром которой служит подогнанная регрессия $(\hat{\beta}_1 + \hat{\beta}_2 x^0)$. Доверительный интервал имеет, очевидно, наименьшую длину, когда x^0 равно наблюденному

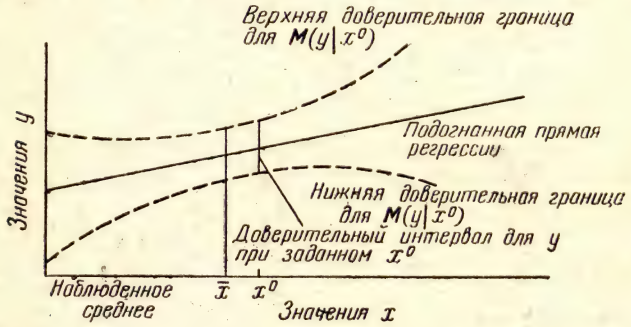


Рис. 28.2. Расположение гиперболических доверительных границ (28.93) для математического ожидания y при простой линейной регрессии.

среднему $\bar{x}$, и его длина возрастает при возрастании $|x^0 - \bar{x}|$. Последнее подтверждает интуитивное представление о том, что наиболее точно мы можем оценивать вблизи «центра» наблюденных значений величины x. Рис. 28.2 иллюстрирует расположение доверительных границ (28.93).

Робинсон (1964) приводит МП-оценки и доверительные интервалы для абсциссы пересечения двух полиномиальных регрессий, а также библиографию относящихся сюда работ.

28.24 Доверительные границы для математического ожидания величины y, рассмотренные в примере 28.4(б) и в более общем

виде в **28.22**, относятся к величине y, соответствующей конкретному значению x^0. На рис. 28.2 каждый конкретный доверительный интервал представляет собой отрезок вертикальной прямой, проходящей через x^0, который заключен между ветвями гиперболы. Предположим теперь, что нам требуется *доверительная область для всей регрессионной прямой*, т. е. область R на плоскости (x, y) (или, в общем случае, в пространстве $(\boldsymbol{x}, y)$), такая, что с вероятностью $1-\alpha$ истинная прямая регрессии $y = \boldsymbol{x}\boldsymbol{\beta}$ содержится в R. Мы видим, что данная задача совершенно отличается от только что рассмотренной. Теперь ищется не интервал, а доверительная область, которая покрывает всю прямую, а не одну точку на прямой. Эта задача в простейшем случае была решена впервые в замечательной работе Уоркинга и Хотеллинга (1929). В нашем изложении мы будем следовать Хоулу (1951).

Доверительная область для регрессионной прямой

28.25 Рассмотрим вначале простой случай примера 28.4 и предположим, что σ^2 известна. Эти ограничения мы ослабим в **28.31—32**. Для удобства будем измерять x_j от их среднего значения, так что $\bar{x} = 0$. При этом анализ, как отмечалось в примере 28.4(а), становится ортогональным. Коэффициенты $\hat{\beta}_1$ и $\hat{\beta}_2$ в этом случае независимы и нормально распределены, а из матрицы рассеяния получаем, что $\mathrm{D}\hat{\beta}_1 = \sigma^2/n$, $\mathrm{D}\hat{\beta}_2 = \sigma^2/\sum x^2$. Таким образом,

$$u = n^{1/2}(\hat{\beta}_1 - \beta_1)/\sigma, \qquad v = (\sum x^2)^{1/2}(\hat{\beta}_2 - \beta_2)/\sigma \tag{28.94}$$

— независимые нормированные нормальные величины.

Пусть $g(u^2, v^2)$ — однозначная четная функция от u и v, и пусть уравнение

$$g(u^2, v^2) = g_{1-\alpha}, \qquad 0 < \alpha < 1, \tag{28.95}$$

определяет семейство замкнутых кривых на плоскости (u, v) таких, что (а) с убыванием $g_{1-\alpha}$ новая кривая содержится внутри кривой, соответствующей большему значению величины $g_{1-\alpha}$; (б) каждая точка, лежащая внутри кривой, принадлежит некоторой другой кривой. Предположим, что неявной зависимости (28.95) между u и v соответствует явная зависимость $u^2 = p(v^2)$ или

$$u = \pm h(v). \tag{28.96}$$

Предположим, далее, что $h'(v) = dh(v)/dv$ существует для всех v и является монотонно убывающей функцией от v, принимающей все действительные значения.

28.26 Из (28.94) видно, что при любом заданном множестве наблюдений, на основании которых подогнана регрессия, истинной регрессионной прямой

$$y = \beta_1 + \beta_2 x \tag{28.97}$$

будут соответствовать значения u и v такие, что

$$\beta_1 + \beta_2 x = \left(\hat{\beta}_1 + \frac{\sigma}{n^{1/2}} u\right) + \left(\hat{\beta}_2 + \frac{\sigma}{\left(\sum x^2\right)^{1/2}} v\right) x. \tag{28.98}$$

Подставляя (28.96) в (28.98), находим два семейства регрессионных прямых, зависящих от параметра v,

$$\left(\hat{\beta}_1 \pm \frac{\sigma}{n^{1/2}} h(v)\right) + \left(\hat{\beta}_2 + \frac{\sigma}{\left(\sum x^2\right)^{1/2}} v\right) x, \tag{28.99}$$

где каждое семейство соответствует определенному знаку в (28.96). Найдем теперь огибающие этих семейств.

Дифференцируя (28.99) по v и приравнивая производную нулю, получаем

$$x = \mp \left(\frac{\sum x^2}{n}\right)^{1/2} h'(v). \tag{28.100}$$

Подставляя (28.100) в (28.99), находим требуемые огибающие

$$(\hat{\beta}_1 + \hat{\beta}_2 x) \pm \frac{\sigma}{n^{1/2}} \{h(v) - v h'(v)\}, \tag{28.101}$$

где функции от v следует заменить функциями от x согласно (28.100). При ограничениях на $h'(v)$, сформулированных после (28.96), две огибающие в (28.101) существуют для всех x, являются однозначными функциями и, кроме того, все прямые каждого семейства лежат только по одну сторону от соответствующей огибающей. Действительно, кривая, отвечающая верхнему знаку в (28.101), всегда лежит выше кривой, отвечающей нижнему знаку в (28.101), а все прямые обоих семейств (28.99) расположены между ними.

28.27 Любая пара значений (u, v), для которой

$$g(u^2, v^2) < g_{1-\alpha}, \tag{28.102}$$

будет соответствовать регрессионной прямой, лежащей между двумя огибающими (28.101), потому что при любом фиксированном v значение $u^2 = \{h(v)\}^2$ уменьшается и, следовательно, функция от v, входящая в постоянный член в (28.99), уменьшается по абсолютной величине, тогда как коэффициент при x не изменяется. Таким образом, если u и v удовлетворяют (28.102), то истинная регрессионная линия лежит между двумя огибающими

(28.101). Выберем теперь $g_{1-\alpha}$ так, чтобы непрерывная случайная величина $g(u^2, v^2)$ удовлетворяла условию

$$\mathbf{P}\{g(u^2, v^2) < g_{1-\alpha}\} = 1 - \alpha. \tag{28.103}$$

Тогда (28.102) выполняется с вероятностью $1-\alpha$ и область R между парой огибающих (28.101) является доверительной областью для истинной прямой регрессии с коэффициентом доверия $1-\alpha$.

28.28 Теперь мы должны рассмотреть вопрос о выборе функции $g(u^2, v^2)$ такой, чтобы при фиксированном значении $1-\alpha$ доверительная область R была в некотором смысле возможно меньшей. Нельзя просто минимизировать площадь области R, так как она всегда бесконечна. Поэтому мы введем весовую функцию $w(x)$ и будем выбирать R так, чтобы минимизировать интеграл

$$I = \int_{-\infty}^{\infty} (y_2 - y_1)\, w(x)\, dx, \tag{28.104}$$

где y_1, y_2 — нижняя и верхняя огибающие (28.101) соответственно, являющиеся границами R, а $\int_{-\infty}^{\infty} w(x)\, dx = 1$. Можно переписать (28.104) в виде

$$I = \mathbf{M}(y_2) - \mathbf{M}(y_1), \tag{28.105}$$

где математические ожидания взяты относительно $w(x)$.

Очевидно, оптимальная область R, найденная с помощью минимизации, будет зависеть от выбранной весовой функции. Полагая $S^2 = \sum x^2/n$, рассмотрим нормальную весовую функцию

$$w(x) = (2\pi S^2)^{-1/2} \exp\left(-\frac{x^2}{2S^2}\right), \tag{28.106}$$

которая особенно удобна, если значения x, считающиеся здесь фиксированными, в действительности выбраны из нормальной совокупности, например, если x и y имеют двумерное нормальное распределение. Подставляя (28.101) и (28.106) в (28.105), получаем

$$I = \frac{2\sigma}{n^{1/2}} [\mathbf{M}\{h(v)\} - \mathbf{M}\{vh'(v)\}]. \tag{28.107}$$

Поскольку $h'(v)$ убывает, из (28.100) имеем

$$dx = -Sh''(v)\, dv, \tag{28.108}$$

так что, преобразуя интегралы в (28.107) к переменной v, находим

$$\left.\begin{aligned} \mathbf{M}\{h\} &= -(2\pi)^{-1/2}\int \hat{h}h''\exp\Big\{-\tfrac{1}{2}(h')^2\Big\}dv, \\ \mathbf{M}(vh') &= -(2\pi)^{-1/2}\int vh'h''\exp\Big\{-\tfrac{1}{2}(h')^2\Big\}dv, \end{aligned}\right\} \tag{28.109}$$

где интегрирования производятся по всей области изменения v. Учитывая четность функции $h(v)$, оба интеграла нужно брать только по положительным v. Подстановка (28.109) в (28.107) дает

$$I = -\frac{4\sigma}{(2\pi n)^{1/2}}\int_0^{v_{\max}} h''(h - vh')\exp\Big\{-\tfrac{1}{2}(h')^2\Big\}dv. \tag{28.110}$$

Эту величину предстоит минимизировать при условии (28.103), которое вследствие независимости, нормальности и симметрии распределений u и v эквивалентно условию

$$(2\pi)^{-1}\int_0^{v_{\max}}\left\{\int_0^{h(v)}\exp\Big(-\tfrac{1}{2}u^2\Big)du\right\}\exp\Big(-\tfrac{1}{2}v^2\Big)dv = \tfrac{1}{4}(1-\alpha). \tag{28.111}$$

Напомним, что мы требуем от функции $h'(v)$, чтобы она была монотонно убывающей по v и принимала все действительные значения.

28.29 Переходя непосредственно к минимизации интеграла I, выберем общий вид для $g(u^2, v^2)$. Здесь «естественно» выбрать семейство эллипсов, которые запишем в форме

$$h(v) = b(a^2 - v^2)^{1/2}. \tag{28.112}$$

Теперь мы должны минимизировать (28.110) по переменной a при условии (28.111). Поскольку из (28.112)

$$\left.\begin{aligned} h'(v) &= -b^2 v/h(v), \\ h''(v) &= -b^2[\{h(v)\}^2 + b^2v^2]/\{h(v)\}^3, \end{aligned}\right\} \tag{28.113}$$

то для выбора a нам надо минимизировать (опуская константу) величину

$$J = b^2\int_0^a \frac{[\{h(v)\}^2 + b^2v^2]^2}{\{h(v)\}^5}\exp\Big[-\tfrac{1}{2}\{b^2v/h(v)\}^2\Big]dv \tag{28.114}$$

при условии

$$\int_0^a \left\{ \int_0^{b(a^2-v^2)^{1/2}} \exp\left(-\tfrac{1}{2}u^2\right) du \right\} \exp\left(-\tfrac{1}{2}v^2\right) dv = k. \quad (28.115)$$

Дифференцируя (28.114) по a и заменяя в этой производной db/da значением, получаемым посредством дифференцирования (28.115), находим после некоторых упрощений

$$\frac{dJ}{da} = b^2 \exp\left(\tfrac{1}{2}b^2\right) \left[\int_0^1 \frac{\exp\left\{-\frac{1}{2}b^2/(1-t^2)\right\} dt}{(1-t^2)^2} - \right.$$

$$\left. - \frac{\int_0^1 \frac{\exp\left\{-\frac{1}{2}b^2/(1-t^2)\right\} dt}{(1-t^2)} \int_0^1 \frac{\exp\left\{-\frac{1}{2}a^2(1-b^2)t^2\right\} dt}{(1-t^2)^{1/2}}}{\int_0^1 (1-t^2)^{1/2} \exp\left\{-\frac{1}{2}a^2(1-b^2)t^2\right\} dt} \right]. \quad (28.116)$$

Для нахождения минимума полагаем $\frac{dJ}{da} = 0$, решаем уравнение относительно a и затем определяем b.

Хоул (1951) проделал это для $1-\alpha = 0{,}95$ и нашел, что эллипс (28.112) имеет полуоси 2,62 и 2,32, мало отличающиеся друг от друга. Если вместо эллипса взять круг, т. е. в (28.112) положить $b = 1$, то его радиус оказывается равным 2,45 при $1-\alpha = 0{,}95$. Хоул показал, что в этом случае значение J отличается от минимального не более чем на один процент.

28.30 Выбор круга в качестве $g(u^2, v^2)$ соответствует первоначальному решению данной задачи Уоркингом и Хотеллингом (1929). Они получили указанное решение, заметив просто, что так как u и v в (28.94) являются независимыми нормированными нормальными величинами, то величина $u^2 + v^2$ имеет χ^2-распределение с 2 степенями свободы, а a^2 $(= g_{1-\alpha}$ в (28.103)) есть $100(1-\alpha)$-процентная точка, получаемая из таблиц χ^2-распределения. Подставляя (28.112) и (28.113) с $b = 1$ в (28.101), находим границы доверительной области

$$(\hat{\beta}_1 + \hat{\beta}_2 x) \pm \frac{\sigma}{n^{1/2}} \left\{ h(v) + \frac{v^2}{h(v)} \right\} =$$

$$= (\hat{\beta}_1 + \hat{\beta}_2 x) \pm \frac{\sigma}{n^{1/2}} [\{h(v)\}^2 + v^2]^{1/2} \left[1 + \frac{v^2}{\{h(v)\}^2} \right]^{1/2} =$$

$$= (\hat{\beta}_1 + \hat{\beta}_2 x) \pm \frac{\sigma}{n^{1/2}} (g_{1-\alpha})^{1/2} [1 + \{h'(v)\}^2]. \quad (28.117)$$

Учитывая (28.100), перепишем (28.117) в виде

$$(\hat{\beta}_1 + \hat{\beta}_2 x) \pm (g_{1-\alpha})^{1/2} \left\{ \frac{\sigma^2}{n} + x^2 \frac{\sigma^2}{\sum x^2} \right\}^{1/2}, \qquad (28.118)$$

где содержимое фигурных скобок равно $\{\mathrm{D}\hat{\beta}_1 + x^2 \mathrm{D}\hat{\beta}_2\}$.

Сравнивая (28.118) с доверительными границами (28.93) для $\mathbf{M}(y|x^0)$, полученными в примере 28.4 (где нужно положить $\bar{x} = 0$, как сделано здесь), мы видим, что с точностью до замены s^2 на σ^2, а $t_{1-\alpha/2}$ на значение $(g_{1-\alpha})^{1/2}$, получаемое из χ^2-распределения, выражения имеют один и тот же вид. Таким образом, доверительные границы (28.118) выглядят так же, как доверительные границы (28.93), изображенные на рис. 28.2 и представляющие собой гиперболу с подгоняемой прямой в качестве диаметра. Как можно было ожидать, при заданном α ветви гиперболы (28.118) быстрее расходятся в стороны, чем их аналоги в (28.93). Это происходит потому, что здесь мы устанавливаем область для целой прямой, тогда как прежде искали границы, заключающие единственное значение на прямой. Например, при $\alpha = 0{,}05$, $t_{1-\alpha/2} = 1{,}96$ (при бесконечном числе степеней свободы, соответствующем известной σ^2), в то время как $g_{1-\alpha}$ для χ^2-распределения с двумя степенями свободы $= 5{,}99$, откуда, извлекая квадратный корень, получаем значение $a = 2{,}45$, приведенное в конце **28.29**.

28.31 Если σ^2 неизвестна, то требуется небольшая модификация рассуждений из **28.25—30**. Определим величину

$$w^2 = (n-2)\, s^2/\sigma^2, \qquad (28.119)$$

так что w^2 есть отношение суммы квадратов остатков от подогнанной регрессии к истинной дисперсии ошибки. Она имеет (см. **28.21**) χ^2-распределение с $n-2$ степенями свободы. Из (28.84) и (28.119) следует, что каждая из статистик

$$u^* = (n-2)^{1/2}\, u/w = n^{1/2} (\hat{\beta}_1 - \beta_1)/s$$

и

$$v^* = (n-2)^{1/2}\, v/w = n^{1/2} (\hat{\beta}_2 - \beta_2)/s$$

подчинена распределению Стьюдента с $n-2$ степенями свободы. Если мы теперь вновь проведем рассуждения из **28.25—30**, подставляя u^* и v^* вместо u и v, то обнаружим, что распределение величины $g(u^{*2}, v^{*2})$ не зависит от параметров β_1, β_2, σ^2. Решение задачи о минимизации взвешенной площади из **28.28—29** становится теперь трудным, и мы перейдем прямо к классическому решению, приведенному в **28.30**.

Будем непосредственно использовать рассуждения Хотеллинга и Уоркинга. Поскольку согласно **28.21** u^2, v^2 и w^2 распределены независимо друг от друга как χ^2-величины с 1, 1 и

$n-2$ степенями свободы соответственно, то отношение $\left(\frac{u^2+v^2}{2}\right)\Big/\left(\frac{w^2}{n-2}\right)=\frac{1}{2}\left(u^{*2}+v^{*2}\right)$ имеет F-распределение с 2 и $n-2$ степенями свободы. Таким образом, если в **28.30** заменить σ на s и положить $g_{1-\alpha}$ равным удвоенной $100(1-\alpha)$-процентной точке этого F-распределения, то из (28.118) мы получим требуемую доверительную область. Как и в **28.30**, обнаруживаем, что границы области всегда расходятся в стороны быстрее, чем границы для $\mathbf{M}(y|x^0)$.

28.32 Нетрудно обобщить наши результаты на случай более чем одного регрессора, и набросок такого обобщения сделан Хоулом (1951). Обобщая результат из **28.31** на k регрессоров, находим, что величина $\left(u^{*2}+\sum_{i=1}^{k} v_i^{*2}\right)\Big/(k+1)$ имеет распределение дисперсионного отношения с $(k+1,\ n-k-1)$ ст. св.

Для полиномиальной регрессии Гафариан (1964) приводит метод получения доверительных областей над любым подмножеством области определения регрессии. В случае прямой линии у него приложены подробные таблицы для области с постоянной шириной.

УПРАЖНЕНИЯ

28.1 Величины x и y имеют двумерное равномерное распределение на плоскости (x, y) в области, ограниченной эллипсом

$$ax^2+2hxy+by^2=c, \quad h\neq 0, \quad h^2<ab; \quad a, b>0.$$

Показать, что регрессия каждой величины относительно другой линейна и что каждая скедастическая кривая есть квадратичная парабола.

28.2 Двумерное распределение величин x и y равномерно на параллелограмме, ограниченном линиями $x=3(y-1)$, $x=3(y+1)$, $x=y+1$, $x=y-1$. Показать, что регрессия y по x линейна, но что регрессия x по y состоит из трех соединенных прямолинейных отрезков.

28.3 Пусть выполняются соотношения (28.59), (28.60), но элементы матрицы $\boldsymbol{X}$ являются нелинейными функциями от r дополнительных параметров $\gamma_1, \ldots, \gamma_r$, т. е. всего имеется $k+r$ параметров. Показать (см. **19.13—16**), что регрессионную модель можно представить в виде

$$y=(\boldsymbol{X}, \boldsymbol{D})\begin{pmatrix}\boldsymbol{\beta}\\ \mathbf{0}\end{pmatrix}+\varepsilon,$$

где $\boldsymbol{D}$ — некоторая матрица $(n\times r)$, выбранная так, чтобы ранг матрицы $(\boldsymbol{X}, \boldsymbol{D})$ равнялся $(k+r)$, а $\mathbf{0}$ — нулевой вектор $(r\times 1)$. Получить отсюда доверительные области для (а) полного набора из $k+r$ параметров, (б) одних только r дополнительных параметров.

(Гальперин 1963; см. также Хартли (1964).)

28.4 Исходя из (28.17) показать, что если маргинальное распределение двумерного распределения имеет форму Грама — Шарлье

$$f=\alpha(x)\{1+a_3H_3+a_4H_4+\ldots\},$$

то регрессия величины y по x есть

$$\mu'_{1x} = \frac{\sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{\varkappa_{r1}}{r!} a_s H_{r+s}(x)}{1 + \sum_{r=3}^{\infty} a_r H_r(x)}.$$

(Уикселл, 1917.)

28.5 Величины x_1, x_2, x_3 имеют трехмерное нормальное распределение. На основании **28.9** показать, что регрессия любой из них относительно двух других линейна.

28.6 Проверить равенство (28.33).

28.7 Пусть при каждом фиксированном y условное распределение величины x нормально. Доказать, что их двумерное распределение должно иметь вид

$$f(x, y) = \exp\{-(a_1 x^2 + a_2 x + a_3)\},$$

где a_i — функции от y. Показать, что если, кроме того, равновероятные контуры функции $f(x, y)$ суть подобные концентрические эллипсы, то f должна быть двумерной нормальной плотностью.

(Бхаттачария, 1943.)

28.8 Показать, что если регрессия x по y линейна, условное распределение x при каждом фиксированном y нормально и гомоскедастично, а маргинальное распределение y нормально, то $f(x, y)$ — двумерная нормальная плотность.

(Бхаттачария, 1943.)

28.9 Показать, что если условные распределения величины x при каждом фиксированном y и величины y при каждом фиксированном x нормальны и одно из них гомоскедастично, то $f(x, y)$ — двумерная нормальная плотность.

(Бхаттачария, 1943.)

28.10 Показать, что если каждая невырожденная линейная функция от x и y имеет нормальное распределение, то $f(x, y)$ — двумерная нормальная плотность.

(Бхаттачария, 1943.)

28.11 Показать, что если регрессии x по y и y по x линейны и условные распределения каждой величины при любом фиксированном значении другой величины нормальны, то функция $f(x, y)$ или является двумерной нормальной плотностью, или может быть записана (при подходящем выборе масштаба и начала координат) в виде

$$f = \exp\{-(x^2 + a^2)(y^2 + b^2)\}$$

(Бхаттачария, 1943.)

28.12 Пусть строится интервальная оценка для β в модели линейной регрессии $y_i = \beta x_i + \varepsilon_i$. Показать, что интервал, основанный на величине Стьюдента

$$t = (b - \beta) \Big/ \left(s^2 \Big/ \sum x_i^2\right)^{1/2},$$

при любой выборке физически короче, чем интервал, построенный с помощью величины

$$u = (\bar{y} - \beta\bar{x}) / (s^2/n)^{1/2}.$$

28.13 Доказать, что если при построении доверительных областей для линии регрессии в **28.28** использовать весовую функцию

$$w(x) = \left(1 + \frac{x}{S^2}\right)^{-3/2}$$

вместо (28.106), то решение Уоркинга — Хотеллинга из **28.30** будет строго оптимальным (минимизирующим площадь) в семействе эллипсов (28.112). (Хоул, 1951.)

28.14 Пусть имеется два различных вектора $\boldsymbol{y}_1$ и $\boldsymbol{y}_2$, каждый из которых связан с одним и тем же множеством регрессоров $\boldsymbol{x}$ в некоторой линейной модели. Показать, что для проверки гипотезы о разности между любыми двумя соответствующими параметрами моделей можно применить метод из **28.21** к разностям $(y_{1i} - y_{2i})$.

(Иэйтс (1939b) рассмотрел также случай, когда регрессоры различны, а векторы $\boldsymbol{y}$ коррелированы.)

28.15 Независимые выборки объемов n_i извлечены из двух регрессионных моделей

$$y = \alpha_i + \beta_i x + \varepsilon, \qquad i = 1, 2,$$

с независимыми нормально распределенными ошибками. В обоих моделях дисперсии ошибок равны σ^2. Пусть b_1, b_2 — отдельные оценки наименьших квадратов для β_1, β_2. Показать, что разность $(b_1 - b_2)$ распределена нормально со средним $(\beta_1 - \beta_2)$ и дисперсией

$$\sigma^2 \left\{ \left(\sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2 \right)^{-1} + \left(\sum_{j=1}^{n_2} (x_j - \bar{x}_2)^2 \right)^{-1} \right\}$$

и что величина

$$t = \{(b_1 - b_2) - (\beta_1 - \beta_2)\} \Bigg/ \left\{ s^2 \left(\frac{1}{\sum\limits_i (x_i - \bar{x}_1)^2} + \frac{1}{\sum\limits_j (x_j - \bar{x}_2)^2} \right) \right\}^{1/2}$$

имеет t-распределение Стьюдента с $n_1 + n_2 - 4$ степенями свободы, где

$$s^2 = \frac{(n_1 - 2)\, s_1^2 + (n_2 - 2)\, s_2^2}{n_1 + n_2 - 4},$$

а s_1^2, s_2^2 — отдельные оценки для σ^2 по двум моделям. Показать, следовательно, что статистику t можно использовать для проверки гипотезы $\beta_1 = \beta_2$ против $\beta_1 \neq \beta_2$.

(См. Фишер, 1922b.)

28.16 Пусть для простой линейной модели $y = \beta_0 + \beta_1 x + \varepsilon$ две независимые выборки объемов m и n имеют средние $(\bar{y}_m, \bar{x}_m)$ и $(\bar{y}_n, \bar{x}_n)$. Показать, что для параметра β_1 величина $b_1 = (\bar{y}_m - \bar{y}_n)/(\bar{x}_m - \bar{x}_n)$ служит несмещенной оценкой, дисперсия которой равна $\sigma^2 \left(\frac{1}{m} + \frac{1}{n}\right) \Big/ (\bar{x}_m - \bar{x}_n)^2$, и что оценка b_1 несостоятельна (когда $m, n \to \infty$, а отношение m/n фиксировано), если две выборки были образованы случайным разбиением одной исходной выборки, содержащей $m + n$ наблюдений.

28.17 Пусть нам даны n наблюдений над моделью $y = \beta_1 x_1 + \beta_2 x_2 + \varepsilon$ с дисперсией ошибки σ^2 и независимая от них несмещенная оценка b_1 для β_1 вместе с несмещенной оценкой s_1^2 ее дисперсии σ_1^2. С целью оценить β_2 рассмотрим регрессию величины $(y - b_1 x_1)$ по x_2. Показать, что оценка $b_2 = \sum (y - b_1 x_1)\, x_2 / \sum x_2^2$ является несмещенной, а ее дисперсия равна

$$\mathrm{D}b_2 = \left(\sigma^2 + \sigma_1^2 r^2 \sum x_1^2\right) / \sum x_2^2,$$

где r — выборочная корреляция между x_1 и x_2. Показать, что если оценка b_1 не дана, то обычная оценка наименьших квадратов для β_2 имеет дисперсию $\sigma^2 / \left\{\sum x_2^2 (1 - r^2)\right\}$, и что, следовательно, использование внешней информации о β_1 увеличивает эффективность оценки для β_2 тогда и только тогда, когда

$$\sigma_1^2 < \frac{\sigma^2}{\sum x_1^2 (1 - r^2)},$$

т. е. если дисперсия оценки b_1 меньше, чем дисперсия обычной оценки наименьших квадратов для β_1. Показать, что несмещенная оценка для $\mathrm{D}b_2$ дается соотношением

$$\hat{V} = \frac{1}{(n-2) \sum x_2^2} \left[\sum (y - b_1 x_1 - b_2 x_2)^2 + s_1^2 \sum x_1^2 \left\{(n-1) r^2 - 1\right\}\right],$$

но что если ошибки распределены нормально, то она распределена не как величина, кратная χ^2.

(Дёрбин, 1953.)

28.18 Обобщим ситуацию упражнения 28.17. Пусть $\boldsymbol{b}_1$ — вектор несмещенных оценок для h параметров $(\beta_1, \beta_2, \ldots, \beta_h)$ с матрицей рассеяния $\boldsymbol{V}_1$, и пусть $\boldsymbol{b}_2$ — не зависящий от него вектор несмещенных оценок для k $(>h)$ параметров $(\beta_1, \beta_2, \ldots, \beta_h, \beta_{h+1}, \ldots, \beta_k)$ с матрицей рассеяния $\boldsymbol{V}_2$. Используя обобщение Эйткина гауссовой теоремы наименьших квадратов (**19.17**), показать, что несмещенные оценки параметров $(\beta_1, \ldots, \beta_k)$ с минимальными дисперсиями, линейные относительно компонент векторов $\boldsymbol{b}_1$ и $\boldsymbol{b}_2$, представляют собой компоненты вектора

$$\boldsymbol{b} = \left\{\left(\boldsymbol{V}_1^{-1}\right)^* + \boldsymbol{V}_2^{-1}\right\}^{-1} \left\{\left(\boldsymbol{V}_1^{-1}\right)^* \boldsymbol{b}_1^* + \boldsymbol{V}_2^{-1} \boldsymbol{b}_2\right\}$$

с матрицей рассеяния

$$\boldsymbol{V}(\boldsymbol{b}) = \left\{\left(\boldsymbol{V}_1^{-1}\right)^* + \boldsymbol{V}_2^{-1}\right\}^{-1},$$

где звездочка обозначает превращение $(h \times 1)$-мерного вектора в $(k \times 1)$-мерный или $(h \times h)$-мерной матрицы в $(k \times k)$-мерную матрицу путем увеличения порядка (исходная матрица помещается в левый верхний угол) и заполнения пустых мест нулями.

Показать, что в частном случае $h = 1$ $\boldsymbol{V}(\boldsymbol{b})$ сводится к

$$\boldsymbol{V}(\boldsymbol{b}) = \sigma^2 \begin{Vmatrix} \sum x_1^2 + \dfrac{\sigma^2}{\sigma_1^2} & \sum x_1 x_2 & \ldots & \sum x_1 x_k \\ \sum x_1 x_2 & \sum x_2^2 & \ldots & \sum x_2 x_k \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \sum x_1 x_k & \sum x_2 x_k & & \sum x_k^2 \end{Vmatrix}^{-1}.$$

что отличается только членом в левом верхнем углу от обычной матрицы рассеяния $\sigma^2(X'X)^{-1}$, полученной по методу наименьших квадратов.

(Дёрбин, 1953.)

28.19 Для подгонки обычной (получаемой по методу наименьших квадратов) регрессии y по x может быть использована простая графическая процедура *без вычислений* в случае, когда значения x равно отстоят друг от друга, скажем на интервал s. Пусть n наблюденных точек на диаграмме разброса (y, x) обозначены $P_1, P_2, \ldots, P_n$ в порядке возрастания x. На P_1P_2 ищем точку Q_2 с координатой x, на $\frac{2}{3}s$ большей, чем у P_1; на Q_2P_3 ищем Q_3 с координатой x, на $\frac{2}{3}s$ большей, чем у Q_2. И так далее одинаковыми шагами, присоединяя каждую точку Q к следующей точке P и находя новую точку Q с увеличенной на $\frac{2}{3}s$ x-координатой, пока последний отрезок $Q_{n-1}P_n$ не даст последнюю точку Q_n. Проведем ту же процедуру в обратном направлении, начиная с отрезка P_nP_{n-1} и определяя, скажем, Q_2' с координатой x, на $\frac{2}{3}s$ меньшей, чем у P_n. И вновь до тех пор, пока не получим точку Q_n' на отрезке $Q_{n-1}'P_1$ с координатой x, на $\frac{2}{3}s$ меньшей, чем у Q_{n-1}'. Тогда Q_nQ_n' есть линия, получаемая по методу наименьших квадратов. Доказать это.

(Асковиц, 1957.)

28.20 Предположим, что мы обратили матрицу $\boldsymbol{A}$, состоящую из сумм квадратов и перекрестных произведений n наблюдений над p величинами. Пусть имеется вектор $\boldsymbol{x}$, содержащий по одному новому наблюдению над каждой величиной. Показать, что матрица, обратная к $\boldsymbol{B} = \boldsymbol{A} + \boldsymbol{xx}'$, представима в виде

$$\boldsymbol{B}^{-1} = \boldsymbol{A}^{-1} - (\boldsymbol{A}^{-1}\boldsymbol{xx}'\boldsymbol{A}^{-1})/(1 + \boldsymbol{x}'\boldsymbol{A}^{-1}\boldsymbol{x}).$$

Вывести отсюда, что для квадратичной формы $\boldsymbol{x}'\boldsymbol{A}^{-1}\boldsymbol{x}$ выполняется соотношение

$$1 + \boldsymbol{x}'\boldsymbol{A}^{-1}\boldsymbol{x} = |\boldsymbol{A} + \boldsymbol{xx}'|/|\boldsymbol{A}|.$$

(См. Бартлетт, 1951.)

28.21 В регрессионной модели

$$y_i = \alpha + \beta x_i + \varepsilon_i, \qquad i = 1, 2, \ldots, n,$$

предположим, что наблюденное среднее $\bar{x} = 0$. Пусть x_0 удовлетворяет уравнению $\alpha + \beta x_0 = 0$. Используя случайную величину $\hat{\alpha} + \hat{\beta}x_0$, получить для квадратичной функции доверительное утверждение вида

$$\mathrm{P}\{Q(x_0) \geqslant 0\} = 1 - \alpha.$$

Вывести отсюда доверительное утверждение для самой величины x_0 и показать, что в зависимости от коэффициентов квадратичной функции оно может поместить x_0

(I) в конечный интервал;
(II) вне конечного интервала;
(III) в бесконечный интервал, состоящий из всей числовой прямой.

(См. Леман, 1959.)

28.22 С целью определения того, какая из моделей

$$y = \beta_0' + \beta_1' x_1 + \varepsilon', \qquad y = \beta_0'' + \beta_2' x_2 + \varepsilon''$$

более эффективна для предсказания y, рассмотреть модель

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i, \qquad i = 1, 2, \ldots, n,$$

с независимыми нормальными ошибками, дисперсия которых σ^2 оценивается через s^2 с $n-2$ степенями свободы. Показать, что статистики

$$z_s = \sum_i (y_i - \bar{y})(x_{si} - \bar{x}_s) \Big/ \left\{ \sum_i (x_{si} - \bar{x}_s)^2 \right\}^{1/2}, \quad s = 1, 2,$$

имеют

$$\mathrm{D}z_1 = \mathrm{D}z_2 = \sigma^2, \qquad \mathrm{cov}\,(z_1, z_2) = \sigma^2 r_{12},$$

где r_{12} — выборочная корреляция между x_1 и x_2. Доказать далее, что распределение разности $(z_1 - z_2)$ в точности нормально со средним $\beta_1' \left\{ \sum_i (x_{1i} - \bar{x}_1)^2 \right\}^{1/2} - \beta_2' \left\{ \sum_i (x_{2i} - \bar{x}_2)^2 \right\}^{1/2}$ и дисперсией $2\sigma^2(1 - r_{12})$. Используя тот факт, что $\sum_i (y_i - \bar{y})^2 - (\beta_s')^2 \sum_i (x_{si} - \bar{x}_s)^2$ есть сумма квадратов отклонений от регрессии y только по x_s, показать, что гипотеза о равенстве этих двух сумм квадратов может быть проверена с помощью статистики $t = (z_1 - z_2)/\{2s^2(1 - r_{12})\}^{1/2}$, имеющей распределение Стьюдента с $n - 3$ степенями свободы.

(Хотеллинг (1940); Хили (1955). Э. Уилльямс (1959) обобщил этот критерий на число моделей, предсказывающих y, большее, чем два.)

28.23 Рассматривая случай, когда в точности $y_j = x_j^k$, $j = 1, 2, \ldots, n$, показать, что если ортогональные полиномы, определенные в (28.73) и (28.74), *ортонормальны* (т. е. $\sum_{i=1}^{n} \varphi_i^2(x_j) = 1$ при всех i), то они удовлетворяют рекуррентному соотношению

$$\varphi_k(x_j) = \frac{1}{b_k} \left\{ x_j^k - \sum_{i=0}^{k-1} \varphi_i(x_j) \sum_{j=1}^{n} x_j^k \varphi_i(x_j) \right\},$$

где нормирующая константа b_k определяется из равенства

$$b_k^2 = \sum_{j=1}^{n} \left\{ x_j^k - \sum_{i=0}^{k-1} \varphi_i(x_j) \sum_{j=1}^{n} x_j^k \varphi_i(x_j) \right\}^2.$$

Далее, делая соответствующие изменения, проверить (28.80) и (28.81).

(Робсон, 1959.)

28.24 Для линейной модели $\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$ показать, что НК-оценки могут быть записаны в виде

$$\hat{\boldsymbol{\beta}}_1 = (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'(\boldsymbol{y} - \boldsymbol{X}_2\hat{\boldsymbol{\beta}}_2), \qquad \hat{\boldsymbol{\beta}}_2 = (\boldsymbol{X}_2'\boldsymbol{D}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{D}\boldsymbol{y},$$

где $\boldsymbol{D} = \boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'$.

Показать, что если сначала оценить β_1 из соотношения

$$y = X_1\beta_1 + \varepsilon^*, \qquad \text{(A)}$$

а затем, используя остатки y_r в (A), как если бы они были некоррелированы, оценить β_2 из соотношения

$$y_r = X_2\beta_2 + \eta,$$

то получатся оценки

$$\beta_1^* = (X_1'X_1)^{-1}X_1'y, \qquad \beta_2^* = (X_2'X_2)^{-1}X_2'Dy,$$

являющиеся смещенными оценками всегда, кроме случая, когда $X_1'X_2 = 0$ или $\beta_2 = 0$. Пусть β_2 — скалярный параметр. Доказать, что

$$\beta_2^* = (1 - R^2)\,\hat{\beta}_2,$$

где R — множественный коэффициент корреляции между величиной x_2 и всеми величинами в X_1.

В случае, когда $y_j = \beta_1 x_{1j} + \beta_2 x_{2j} + \varepsilon_j$, показать, что среднеквадратичные ошибки смещенных двухэтапных оценок β_1^*, β_2^* меньше, чем дисперсии несмещенных НК-оценок $\hat{\beta}_1$, $\hat{\beta}_2$, если $\beta_2^2/V(\hat{\beta}_2) < 1$.

(См. Френд и др. (1961), Голдбергер и Джокемс (1961), Голдбергер (1961), Зискинд (1963) и Уоллес (1964).)

ГЛАВА 29

ФУНКЦИОНАЛЬНАЯ И СТРУКТУРНАЯ ЗАВИСИМОСТЬ

Функциональная зависимость между математическими переменными

29.1 В естественных, а отчасти и в социальных науках довольно привычны модели, в которых некоторые математические (не случайные) переменные функционально связаны. Хорошо известным примером является закон Бойля, утверждающий, что при постоянной температуре давление P и объем V заданного количества газа связаны уравнением

$$PV = \text{const.} \tag{29.1}$$

Соотношение (29.1) может, однако, не выполняться, когда газ находится в состоянии, близком к сжижению, а также, возможно, при некоторых других P и V. Если изучается связь между объемом и давлением при так называемом адиабатическом расширении, когда недостаточно времени для теплообмена с окружающей средой, то вместо (29.1) приходится использовать соотношение

$$PV^{\gamma} = \text{const}, \tag{29.2}$$

где γ — константа, которую, возможно, надо оценить. Более того, в некоторых случаях желательно принимать в расчет также температуру T, что приводит к следующему обобщению (29.1):

$$PVT^{-1} = \text{const}.$$

В общем случае мы имеем k переменных $X_1, \ldots, X_k$, связанных между собой с помощью p функциональных соотношений

$$f_j(X_1, \ldots, X_k; \alpha_1, \ldots, \alpha_l) = 0, \quad j = 1, 2, \ldots, p, \tag{29.3}$$

в которые входят l параметров α_r, $r = 1, 2, \ldots, l$. Обычно требуется оценить α_r, исходя из наблюдений, а иногда также определить форму функциональной зависимости f_j. Последняя задача чаще всего возникает в ситуациях, когда вид f_j не удается определить полностью ни из теоретических соображений, ни из

предшествующего опыта. Если бы значения X наблюдались без ошибок, то вообще не возникло бы никакой статистической проблемы: мы имели бы набор значений, удовлетворяющих (29.3), и оставалось бы только решить эту систему, что является чисто математической задачей. Однако обычно на наши измерения влияют ошибки эксперимента или наблюдений. Мы наблюдаем не «истинное» значение X, а значение, отличающееся от него на некоторую случайную величину. Таким образом, задача заключается в оценивании параметров α_r (и, возможно, в определении вида зависимости f_j) исходя из данных, которые, по крайней мере частично, являются выборками из распределений ошибок. В этом случае статистический характер проблемы не подлежит сомнению.

29.2 На наш взгляд, при изложении данного предмета, в котором в прошлом было немало путаницы, чрезвычайно важно придерживаться ясной терминологии и обозначений. В этой главе мы будем обозначать математические переменные большими латинскими буквами. Как обычно, параметры будут обозначаться малыми греческими буквами (особенно часто буквами α и β), а случайные величины — малыми латинскими буквами или, в случае оценок максимального правдоподобия, буквой, обозначающей параметр, с крышечкой наверху, например: $\hat{\alpha}$. Случайные ошибки будут обозначаться другими малыми греческими буквами, чаще всего буквами δ и ε, а наблюдаемые случайные величины, соответствующие ненаблюдаемым переменным, — «соответствующими» *) греческими буквами, например, ξ будет соответствовать X. Единственным источником путаницы в нашей системе обозначений может быть то, что греческие буквы выполняют сразу три роли (параметры, ошибки, наблюдаемые величины), однако для разных целей нами используются разные группы букв, причем соблюдается следующее простое правило: любая греческая буква, «соответствующая» большой латинской букве, обозначает наблюдаемую случайную величину, соответствующую этой математической переменной, тогда как все остальные греческие буквы обозначают ненаблюдаемые величины (параметры или ошибки).

29.3 Начнем с самого простого случая. Пусть две математические переменные X и Y связаны линейным соотношением

$$Y = \alpha_0 + \alpha_1 X \tag{29.4}$$

*) Упомянутое «соответствие» между латинскими и греческими буквами основывается не столько на сходстве алфавитов, сколько на сходстве в написании и произношении. В любом случае было бы логичней использовать обычные малые латинские буквы, т. е. обозначать x наблюдаемую величину, соответствующую математической переменной X, однако тогда появилась бы опасность путаницы в индексах и, кроме того, обозначение x нам понадобится для других целей (см. **29.6**).

и требуется оценить параметры α_0, α_1. Мы не можем наблюдать значения X и Y, но можем наблюдать значения случайных величин ξ и η, определенных следующим образом:

$$\left.\begin{aligned} \xi_i &= X_i + \delta_i, \\ \eta_i &= Y_i + \varepsilon_i, \end{aligned}\right\} \quad i = 1, 2, \ldots, n. \tag{29.5}$$

Индексы в (29.5) существенны: наблюдения распределены около «истинного» значения в соответствии с распределением «ошибки», которое может зависеть от i. Например, ошибки могут быть больше для больших значений X, что может быть выражено увеличением дисперсии, соответствующей ошибке случайной величины.

В рассматриваемом простейшем случае будет, однако, предполагаться, что δ_i распределены одинаково, в частности имеют одно и то же среднее (которое без потери общности можно считать равным нулю) и дисперсию для всех X_i. Аналогичные предположения считаются выполненными для ε и Y. Предполагается также, что ошибки δ и ε некоррелированы между собой и друг с другом. Пока мы не будем предполагать δ и ε нормально распределенными. Итак, наша модель описывается соотношениями (29.4) и (29.5), в которых

$$\left.\begin{aligned} \mathsf{M}(\delta_i) = \mathsf{M}(\varepsilon_i) &= 0, \\ \mathsf{D}\delta_i = \sigma_\delta^2, \quad \mathsf{D}\varepsilon_i &= \sigma_\varepsilon^2 \quad \text{для всех} \quad i, \\ \operatorname{cov}(\delta_i, \delta_j) = \operatorname{cov}(\varepsilon_i, \varepsilon_j) &= 0, \quad i \neq j, \\ \operatorname{cov}(\delta_i, \varepsilon_j) &= 0 \quad \text{для всех} \quad i, j. \end{aligned}\right\} \tag{29.6}$$

Относительно средних δ_i (и ε_i) предполагается фактически их равенство между собой, и их можно положить равными нулю, включая средние μ_δ и μ_ε ошибок δ и ε в α_0, поскольку ясно, что не имеется возможности отделить μ_δ и μ_ε от α_0.

Учитывая (29.6), можно, не вызывая неоднозначного понимания, записать рассматриваемую модель в виде

$$\xi = X + \delta, \qquad \eta = Y + \varepsilon. \tag{29.7}$$

29.4 На первый взгляд оценивание параметров в (29.4) кажется задачей, похожей на задачу регрессионного анализа, и это сходство часто было источником недоразумений. В регрессионных задачах мы, однако, интересуемся зависимостью среднего значения η (которое равно Y) от переменной X, не подверженной ошибкам; в этом случае ошибка δ тождественно равна нулю и, следовательно, $\sigma_\delta^2 = 0$. Таким образом, регрессионная модель есть частный случай рассматриваемой. Кроме того (хотя это уже относится не к формальной постановке задачи, а к ее происхождению), случайные изменения зависимой переменной

в регрессионном анализе не обязательно связаны только с ошибкой измерений. Они могут полностью или частично порождаться самой структурой связи между переменными. Так, между весом тела и ростом существует внутренняя связь, совсем не зависящая от ошибок измерений.

Легко убедиться, что наличие ошибок как в X, так и в Y ставит проблемы, совершенно отличные от тех, которые возникают в задачах регрессионного анализа. Подставляя X и Y из (29.7) в (29.4), мы получаем соотношение

$$\eta = \alpha_0 + \alpha_1\xi + (\varepsilon - \alpha_1\delta), \tag{29.8}$$

которое не совпадает с обычной постановкой задачи в регрессионном анализе, так как случайная величина ξ коррелирована с ошибкой $(\varepsilon - \alpha_1\delta)$. Действительно, из (29.6) и (29.7) получаем

$$\operatorname{cov}(\xi, \varepsilon - \alpha_1\delta) = \mathbf{M}\{\xi(\varepsilon - \alpha_1\delta)\} = \mathbf{M}\{(X+\delta)(\varepsilon - \alpha_1\delta)\} = -\alpha_1\sigma_\delta^2. \tag{29.9}$$

(29.9) обращается в нуль либо при $\sigma_\delta^2 = 0$, что совпадает с регрессионной ситуацией, либо в тривиальном случае $\alpha_1 = 0$.

Уравнение (29.8) называется *структурным соотношением* между наблюдаемыми случайными величинами ξ и η. Это структурное соотношение порождается *функциональным соотношением* между математическими переменными X и Y.

29.5 В регрессионном анализе можно было выбирать значения независимой переменной X произвольно, например, через равные промежутки в области ее изменения. Однако эти значения могли быть получены и в результате случайного выбора, т. е. n пар наблюдений могли быть случайно выбраны из двумерного распределения с целью изучения регрессии одной случайной величины по другой. (Эти варианты регрессионных моделей уже рассматривались в **26.24**, **27.29**.) В данной модели значения X также могут быть получены либо в результате некоторого случайного процесса, либо путем целенаправленных измерений в выбранных точках, однако в любом случае X останется ненаблюдаемой величиной вследствие ошибок измерений. Теперь мы переходим к изучению ситуации, когда X и Y являются случайными величинами, так что функциональное соотношение (29.4) превращается в структурное соотношение между ненаблюдаемыми величинами.

Структурные соотношения между случайными величинами

29.6 Предположим, что X и Y сами являются случайными величинами (которые в соответствии с принятыми соглашениями будут обозначаться x и y) и что, как и прежде, выполнены (29.4), (29.5) и (29.6). Отсюда можно снова вывести (29.8),

однако (29.9) уже не будет выполнено, если не сделать дополнительных предположений, поскольку в (29.9) величина X рассматривалась как константа. В действительности мы получим

$$\operatorname{cov}(\xi, \varepsilon - \alpha_1\delta) = \mathbf{M}\{(x+\delta)(\varepsilon - \alpha_1\delta)\} = \mathbf{M}(x\varepsilon) - \alpha_1\mathbf{M}(x\delta) - \alpha_1\sigma_\delta^2. \tag{29.10}$$

Сделав, однако, дополнительные предположения (два относительно x и два относительно y)

$$\operatorname{cov}(x, \delta) = \operatorname{cov}(x, \varepsilon) = \operatorname{cov}(y, \delta) = \operatorname{cov}(y, \varepsilon) = 0, \tag{29.11}$$

мы сведем (29.10) к (29.9).

Таким образом, рассматриваемая модель записывается в виде

$$\left.\begin{aligned} \xi_i &= x_i + \delta_i, \\ \eta_i &= y_i + \varepsilon_i, \end{aligned}\right\} \tag{29.12}$$

$$y_i = \alpha_0 + \alpha_1 x_i, \tag{29.13}$$

причем выполнены также (29.6) и (29.11), так что, как и раньше, справедливо (29.8). В итоге мы заменили функциональное соотношение (29.4) между математическими переменными структурным соотношением (29.13), выражающим точную линейную связь между двумя ненаблюдаемыми случайными величинами x и y. Предыдущая модель получается из этой как частный случай, когда x_i вырождаются в константы X_i. Как и прежде, соотношение (29.8) между наблюдаемыми переменными ξ и η является структурным, но, кроме этого, мы имеем структурное соотношение, так сказать, внутри самой модели.

Модели, описываемые структурными соотношениями, в основном находят применение в социальных науках, особенно в эконометрике. Мы еще возвратимся к этой теме при изложении многомерного анализа в третьем томе. Упомянем только в целях иллюстрации такой пример. Пусть выполнена гипотеза (29.13) о линейной зависимости между количеством проданного товара (y) и его ценой (x), причем x и y считаются случайными величинами. Пусть, кроме того, эти величины наблюдаются с ошибками, т. е. выполняется (29.12). Тогда мы находимся в ситуации, описываемой структурным соотношением. Существенно, однако, что кроме ошибок, возникающих при наблюдении x и y, имеют место случайные изменения этих переменных, обусловленные самой их природой.

29.7 Исследователей часто ставила в тупик следующая ситуация. Рассматривая линейную зависимость между переменными, они не могли допустить существования двух прямых линий регрессии, но в то же время на ранней стадии развития

предмета (а, возможно, иногда и сейчас) им всегда приходилось иметь дело с парой прямых. Однако из наших рассуждений ясно, что линия регрессии и не предназначена для выражения функциональной зависимости между математическими переменными или структурной зависимости между случайными переменными: она либо выражает свойство двумерного распределения, либо, когда независимая переменная не подвержена ошибке, выражает связь между независимой переменной и средним значением зависимой переменной. Методы этой главы, которые, как можно видеть из приводимых ссылок, были главным образом развиты в течение последних двух десятилетий, дают в распоряжение исследователей математические модели, более пригодные для их нужд.

29.8 Интересно выяснить, почему для оценивания α_0 и α_1 в (29.8) нельзя использовать основанный на методе наименьших квадратов регрессионный анализ. Имея n пар наблюденных значений (ξ_i, η_i), $i = 1, 2, \ldots, n$, можно получить, усредняя (29.8) по этим значениям, соотношение

$$\bar{\eta} = \alpha_0 + \alpha_1 \bar{\xi} + \frac{1}{n} \sum (\varepsilon - \alpha_1 \delta). \tag{29.14}$$

Поскольку последнее слагаемое в правой части имеет нулевое математическое ожидание, то мы получаем для оценивания уравнение

$$\bar{\eta} = \alpha_0 + \alpha_1 \bar{\xi}, \tag{29.15}$$

которое является несмещенным в том смысле, что обе его части имеют одинаковое математическое ожидание. Если за начала отсчетов приняты выборочные средние $\bar{\xi}$, $\bar{\eta}$, то для α_0 получается оценка

$$a_0 = 0. \tag{29.16}$$

Аналогично, умножая (29.8) на ξ, находим после усреднения

$$\frac{1}{n} \sum \eta\xi = \frac{a_1}{n} \sum \xi^2 + \frac{1}{n} \sum \xi (\varepsilon - \alpha_1 \delta), \tag{29.17}$$

где a_1 — оценка для α_1. Второе слагаемое в правой части (29.17) не обращается в нуль даже при $n \to \infty$, а стремится к величине $\operatorname{cov}\{\xi, \varepsilon - \alpha_1 \delta\}$, кратной σ_δ^2 в силу (29.9). Таким образом, чтобы оценить α_1, мы должны по крайней мере при использовании этого метода знать σ_δ^2. Далее мы увидим, что дисперсии ошибок действительно играют существенную роль в оценивании α_1.

Оценивание структурных соотношений по методу МП

29.9 Если сделать дополнительное предположение, что все пары наблюдений (ξ_i, η_i) имеют одно и то же совместное нормальное распределение, то мы сможем использовать метод МП для оценивания параметров модели структурной зависимости, определяемой соотношениями (29.6) и (29.11)—(29.13). (Для совместной нормальности достаточно, чтобы все x_i имели одно и то же нормальное распределение, и то же самое было бы верно для y_i, δ_i и ε_i, а в случае, когда x, y вырождаются в константы X, Y, для совместной нормальности ξ и η достаточно двумерной нормальности δ и ε.) Тогда в силу (29.6) и (29.11)—(29.13) мы получим следующие выражения для моментов:

$$\left.\begin{aligned} \mathsf{M}\xi &= \mathsf{M}x = \mu, \\ \mathsf{M}\eta &= \mathsf{M}y = \alpha_0 + \alpha_1\mu, \\ \mathsf{D}\xi &= \mathsf{D}x + \sigma_\delta^2 = \sigma_x^2 + \sigma_\delta^2, \\ \mathsf{D}\eta &= \mathsf{D}y + \sigma_\varepsilon^2 = \alpha_1^2\sigma_x^2 + \sigma_\varepsilon^2, \\ \operatorname{cov}(\xi, \eta) &= \operatorname{cov}(x, y) = \alpha_1\sigma_x^2. \end{aligned}\right\} \qquad (29.18)$$

Следует, в частности, заметить, что в (29.18) структурные переменные x_i (а также y_i) имеют одинаковые средние. Как будет видно, это имеет важное значение при использовании метода МП. Однако это означает, что результаты, которые мы собираемся получить для структурных соотношений, не будут иметь почти никакой ценности в случае функциональных соотношений, поскольку они применимы, только когда X_i (константы, в которые вырождаются x_i при $\sigma_x^2 = 0$) принимают одно и то же значение μ для всех i. См. **29.13** ниже.

29.10 Как следует из (16.47) и из примеров 18.14, 18.15, выборочные средние, дисперсии и ковариация совместно достаточны для пяти параметров двумерного нормального распределения, являясь для них МП-оценками. Поэтому, если s_η^2, s_ξ^2 (>0) — выборочные дисперсии и $s_{\xi\eta}$ — выборочная ковариация, то решения системы уравнений

$$\left.\begin{aligned} &\text{(а)} && \mu = \bar{\xi}, \\ &\text{(б)} && \alpha_0 + \alpha_1\mu = \bar{\eta}, \\ &\text{(в)} && \sigma_x^2 + \sigma_\delta^2 = s_\xi^2, \\ &\text{(г)} && \alpha_1^2\sigma_x^2 + \sigma_\varepsilon^2 = s_\eta^2, \\ &\text{(д)} && \alpha_1\sigma_x^2 = s_{\xi\eta} \end{aligned}\right\} \qquad (29.19)$$

относительно тех из шести параметров μ, α_0, α_1, σ_x^2, σ_δ^2 и σ_ε^2, которые неизвестны, будут МП-оценками для этих параметров при условии, что они дают допустимые значения. Поскольку μ, α_0 и α_1 априори никак не ограничены, то требуется только, чтобы решения для σ_x^2, σ_δ^2 и σ_ε^2 были неотрицательны. Согласно (29.19) (в)—(д) это приводит к следующим ограничениям:

$$\left.\begin{array}{ll}
\text{для } \hat{\sigma}_\delta^2 \geqslant 0 & \text{(i)} \quad \alpha_1 s_\xi^2 \geqslant s_{\xi\eta}, \\
\text{для } \hat{\sigma}_\varepsilon^2 \geqslant 0 & \text{(ii)} \quad s_\eta^2 \geqslant \alpha_1 s_{\xi\eta}, \\
\text{для } \hat{\sigma}_x^2 \geqslant 0 & \left\{\begin{array}{l}
\text{(iii)} \quad s_\xi^2 - \sigma_\delta^2 \geqslant 0, \\
\text{(iv)} \quad s_\eta^2 - \sigma_\varepsilon^2 \geqslant 0, \\
\text{(v) если } \sigma_x^2 > 0, \text{ то } \operatorname{sign} \alpha_1 = \operatorname{sign} s_{\xi\eta}, \\
\text{(vi) если } \sigma_x^2 = 0, \text{ то } \alpha_1 \text{ не определено.}
\end{array}\right.
\end{array}\right\} \quad (29.20)$$

Если ограничения (29.20) не выполнены, то решения (29.19) не будут МП-оценками для нашей задачи. В этом случае МП-оценки находятся путем прямой максимизации ФП. Как показывает (29.18), при этом (29.20) (vi) остается в силе.

Из (29.19) (в)—(д) получаем

$$\alpha_1 (s_\xi^2 - \sigma_\delta^2) = s_{\xi\eta}, \qquad s_\eta^2 - \sigma_\varepsilon^2 = \alpha_1 s_{\xi\eta}, \tag{29.21}$$

откуда, делая равными коэффициенты при α_1, находим

$$\alpha_1 s_{\xi\eta} (s_\xi^2 - \sigma_\sigma^2) = s_{\xi\eta}^2 = (s_\eta^2 - \sigma_\varepsilon^2)(s_\xi^2 - \sigma_\delta^2). \tag{29.22}$$

Если α_1, $s_{\xi\eta} \neq 0$, то из (29.21) следует неравенство

$$\frac{|s_{\xi\eta}|}{s_\xi^2} \leqslant |\alpha_1| \leqslant \frac{s_\eta^2}{|s_{\xi\eta}|}, \tag{29.23}$$

показывающее, что абсолютная величина МП-оценки наклона структурной прямой заключена между коэффициентом регрессии η по ξ, полученным по методу НК, и значением, обратным коэффициенту регрессии ξ по η. Согласно (29.19) (а)—(б) соответствующая структурная прямая лежит между двумя линиями регрессиями, что кажется естественным с интуитивной точки зрения.

29.11 Независимо от того, производится оценивание по методу МП с помощью (29.19) или путем прямой максимизации ФП, уравнения (29.19) (а)—(б) всегда дают МП-оценки $\hat{\mu}$ и $\hat{\alpha}_0$, если оценка $\hat{\alpha}_1$ уже найдена. При использовании (29.19) уравнения (в)—(д) должны быть разрешены относительно α_1, однако этого нельзя сделать без дополнительных предположений, так как рассматриваемые три уравнения содержат четыре неизвестных.

Причину этой трудности легко обнаружить. Возвращаясь (29.18), мы видим, что изменение истинного значения α_1 может не менять значения пяти приведенных там моментов. Пусть, например, μ и α_1 положительны. Тогда любое увеличение α_1 может быть компенсировано (а) уменьшением α_0 в $\mathsf{M}\eta$, (б) уменьшением σ_x^2 в $\operatorname{cov}(\xi, \eta)$ и (в) соответствующим изменением σ_ε^2 в $\mathsf{D}\eta$. (Читатель, возможно, захочет проверить это на численном примере.) Это означает, что параметр α_1 невозможно оценить по самой природе задачи, как бы ни был велик объем выборки. По этой причине он называется *неидентифицируемым*. В действительности из шести параметров идентифицируемым является только μ. Нам нежелательно предполагать известными α_0 и α_1, поскольку их оценивание является нашей главной целью, или σ_x^2, поскольку x — ненаблюдаемая величина. Параметр μ идентифицируем, и здесь положение нельзя улучшить. Очевидно, мы должны сделать предположения относительно дисперсий ошибок.

29.12 *Случай* 1: σ_δ^2 *известна.*

Если $s_\xi^2 > \sigma_\delta^2$, то из (29.21) сразу получаем

$$\hat{\alpha}_1 = s_{\xi\eta}/(s_\xi^2 - \sigma_\delta^2), \tag{29.24}$$

откуда следует, что (29.20) (iii)—(v) выполняются. Однако для того, чтобы были выполнены все ограничения в (29.20) и оценка (29.24) была МП-оценкой, необходимо добавить ограничение (29.20) (ii) в форме $s_\eta^2 \geqslant s_{\xi\eta}^2/(s_\xi^2 - \sigma_\delta^2)$. Если эти ограничения не выполнены, то (29.19) не дает МП-оценки, которая в этом случае равна (см. упражнение 29.18)

$$\hat{\alpha}_1 = s_\eta^2/s_{\xi\eta}. \tag{29.25}$$

Заметим, что если $\sigma_\delta^2 \to 0$, то $|\alpha_1|$ в (29.24) стремится к своей нижней границе в (29.23), тогда как (29.25) дает верхнюю границу.

Случай 2: σ_ε^2 *известна.*

Из (29.21) получаем

$$\hat{\alpha}_1 = (s_\eta^2 - \sigma_\varepsilon^2)/s_{\xi\eta}, \qquad s_{\xi\eta} \neq 0. \tag{29.26}$$

Далее, как и в случае 1, если выполнены условия $s_\eta^2 > \sigma_\varepsilon^2$, $s_\xi^2 \geqslant s_{\xi\eta}^2/(s_\eta^2 - \sigma_\varepsilon^2)$, то будут выполнены все ограничения в (29.20) и (29.26) будет МП-оценкой. Если же эти условия не будут выполнены, то мы получим следующую МП-оценку, аналогичную (29.25):

Последнее утверждение предыдущего случая (с очевидными изменениями) справедливо и здесь.

Случай 3: $\sigma_\varepsilon^2/\sigma_\delta^2$ *известно.*

Это классический метод решения задачи идентифицируемости. Полагая $\sigma_\varepsilon^2 = \lambda\sigma_\delta^2$ и исключая σ_δ^2 из уравнений (29.21), получаем

$$\alpha_1^2 s_{\xi\eta} + \alpha_1(\lambda s_\xi^2 - s_\eta^2) - \lambda s_{\xi\eta} = 0. \tag{29.27}$$

Если $s_{\xi\eta} \neq 0$ (в случае $s_{\xi\eta} = 0$ мы получаем $\hat{\alpha}_1 = 0$, если $s_\eta^2/s_\xi^2 \neq \lambda$; если же $s_\eta^2/s_\xi^2 = \lambda$, то параметр α_1 не определен и $\hat{\sigma}_x^2 = 0$ — см. (29.20) (vi)), то это квадратное уравнение будет иметь ненулевые корни

$$\frac{(s_\eta^2 - \lambda s_\xi^2) \pm \{(s_\eta^2 - \lambda s_\xi^2)^2 + 4\lambda s_{\xi\eta}^2\}^{1/2}}{2s_{\xi\eta}} \equiv \frac{N}{2s_{\xi\eta}}. \tag{29.28}$$

Поэтому в силу (29.19) (д) $\hat{\sigma}_x^2 = s_{\xi\eta}/\alpha_1 = 2s_{\xi\eta}^2/N$, так что для выполнения неравенства $\hat{\sigma}_x^2 \geqslant 0$ N должно быть неотрицательным и, следовательно, надо взять корень со знаком плюс в числителе. Таким образом,

$$\hat{\alpha}_1 = \frac{(s_\eta^2 - \lambda s_\xi^2) + \{(s_\eta^2 - \lambda s_\xi^2)^2 + 4\lambda s_{\xi\eta}^2\}^{1/2}}{2s_{\xi\eta}} \equiv \frac{N_+}{2s_{\xi\eta}} \tag{29.29}$$

и $\hat{\sigma}_x^2 > 0$. Остается только проверить, выполняются ли (29.20) (i) или (ii), поскольку теперь $\hat{\sigma}_\varepsilon^2 = \lambda\hat{\sigma}_\delta^2$. Для выполнения (29.20) (ii) требуется, чтобы $2s_\eta^2 \geqslant N_+$. Это неравенство можно получить, заменяя величину $s_{\xi\eta}^2$ ее верхней границей $s_\xi^2 s_\eta^2$. Таким образом, (29.29) дает МП-оценку.

Случай 4: σ_δ^2 *и* σ_ε^2 *известны.*

В этом случае в (29.19) (в)—(д) остается только два неизвестных (α_1 и σ_x^2), и мы можем получить обе оценки (29.24) и (29.26), которые, однако, несовместимы. Следовательно, с помощью (29.19) нельзя найти МП-оценку, и надо, следуя Бёрчу (1964а), максимизировать непосредственно ФП.

Используя последние три равенства (29.18), получаем

$$-\frac{2}{n}\log L = \log|V| + |V|^{-1}\{s_\xi^2(\alpha_1^2\sigma_x^2 + \sigma_\varepsilon^2) - 2s_{\xi\eta}\alpha_1\sigma_x^2 + s_\eta^2(\sigma_x^2 + \sigma_\delta^2)\}, \tag{29.30}$$

где

$$|V| = (\sigma_x^2 + \sigma_\delta^2)(\alpha_1\sigma_x^2 + \sigma_\varepsilon^2) - (\alpha_1\sigma_x^2)^2 = \alpha_1^2\sigma_x^2\sigma_\delta^2 + \sigma_\varepsilon^2(\sigma_x^2 + \sigma_\delta^2).$$

Для простоты исключим с помощью нормировки известные константы σ_δ и σ_ε из (29.30), измеряя x и ξ в единицах σ_δ, а y и η — в единицах σ_ε. Тогда (29.30) примет вид

$$G(u, v) = -\frac{n}{2}\left\{\log(1+u^2+v^2) + \frac{s_\xi^2(1+v^2) - 2s_{\xi\eta}uv + s_\eta^2(1+u^2)}{1+u^2+v^2}\right\},$$

где $u^2 = \sigma_x^2$, $v^2 = \alpha_1^2 u^2$. Функция $G(u, v)$ дифференцируема и стремится к $-\infty$, когда $(u^2+v^2)\to\infty$, так что она принимает максимальное значение в стационарной точке, получаемой приравниванием к нулю производных $\frac{\partial G}{\partial u}$ и $\frac{\partial G}{\partial v}$. Мы, следовательно, получаем

$$\left.\begin{aligned}(1+u^2)\frac{\partial G}{\partial u} + uv\frac{\partial G}{\partial v} &= n\left(\frac{s_\xi^2 u + s_{\xi\eta}v}{1+u^2+v^2} - u\right) = 0,\\ uv\frac{\partial G}{\partial u} + (1+v^2)\frac{\partial G}{\partial v} &= n\left(\frac{s_{\xi\eta}u + s_\xi^2 v}{1+u^2+v^2} - v\right) = 0,\end{aligned}\right\} \tag{29.31}$$

Исключая из этих уравнений $(1+u^2+v^2)$, находим

$$u^2\{\alpha_1^2 s_{\xi\eta} + \alpha_1(s_\xi^2 - s_\eta^2) - s_{\xi\eta}\} = 0.$$

Таким образом, либо $u^2 = 0$ (и, следовательно, $v^2 = 0$), либо равно нулю выражение в фигурных скобках. В последнем случае мы вновь возвращаемся к (29.27) (вспомним, что вследствие нормировки $\lambda = 1$). Отказываясь от нормировки, мы приходим к выводу, что (29.29) дает решение для α_1. Из (29.31) получаем

$$s_\xi^2 + s_{\xi\eta}\frac{v}{u} - (1+u^2+v^2) = 0,$$

откуда, учитывая равенство $v^2 = \alpha_1^2 u^2$, находим

$$\sigma_x^2 = u^2 = (s_\xi^2 + \alpha_1 s_{\xi\eta} - 1)/(1+\alpha_1^2). \tag{29.32}$$

Значение, принимаемое функцией G в стационарной точке $(\pm u, v)$ и даваемое формулами (29.29) и (29.32), будет максимальным только в случае, если таковым не является значение в точке $u = v = 0$. Из рассмотрения второй производной следует, что для этого требуется выполнение по крайней мере одного из условий $(s_\xi^2 - 1) > 0$, $(s_\eta^2 - 1) > 0$ или $s_{\xi\eta}^2 > (1 - s_\xi^2)\times$ $\times(1 - s_\eta^2)$. Снова отказываясь от нормировки, мы приходим к заключению, что (29.29) и (29.32) дают МП-оценки для (α_1, σ_x^2), когда выполнено по крайней мере одно из условий $s_\xi^2 > \sigma_\delta^2$, $s_\eta^2 > \sigma_\varepsilon^2$ или $s_{\xi\eta}^2 > (\sigma_\delta^2 - s_\xi^2)(\sigma_\varepsilon^2 - s_\eta^2)$. Если же ни одно из них не выполнено (что практически маловероятно), то максимум будет

находиться в точке $u = v = 0$, и тогда $\hat{\sigma}_x^2 = 0$, а параметр α_1 не определен (см. условие (29.20) (vi) и следующее за ним обсуждение).

Проблема идентифицируемости, поставленная в **29.11**, не возникает, если имеются повторные наблюдения, т. е. если имеется r_i наблюдений ξ_{ij} $(j = 1, 2, \ldots, r_i)$, соответствующих истинному значению x_i, и s_i наблюдений η_{ik} $(k = 1, 2, \ldots, s_i)$, соответствующих y_i, причем по крайней мере одно r_i и одно s_i больше единицы. Введем обозначения

$$\xi_{i\cdot} = \sum_{j=1}^{r_i} \xi_{ij}/r_i, \qquad \eta_{i\cdot} = \sum_{k=1}^{s_i} \eta_{ik}/s_i,$$

$$R = \sum_{i=1}^{n} r_i, \qquad S = \sum_{i=1}^{n} s_i,$$

$$\xi_{\cdot\cdot} = \sum_{i=1}^{n} r_i \xi_{i\cdot}/R, \qquad \eta_{\cdot\cdot} = \sum_{i=1}^{n} s_i \eta_{i\cdot}/S.$$

Легко показать, что

$$\hat{\sigma}_\delta^2 = \frac{1}{R-n} \sum_i \sum_j (\xi_{ij} - \xi_{i\cdot})^2$$

и

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{S-n} \sum_i \sum_k (\eta_{ik} - \eta_{i\cdot})^2$$

являются несмещенными оценками дисперсий ошибок и что $\hat{\lambda} = \hat{\sigma}_\varepsilon^2/\hat{\sigma}_\delta^2$ — состоятельная (хотя и смещенная) оценка их отношения. Теперь любая оценка для α_1 (случаи 1—4), если подставить в нее соответствующую оценку из числа приведенных выше, будет состоятельной. Мадански (1959) рассмотрел различные оценки, полученные подобными методами, которые, по существу, представляют простое приложение идей дисперсионного анализа (том 3). Дорфф и Гёрлэнд (1961a) показали, что наименьшая асимптотическая дисперсия при постоянном значении отношения s_i/r_i соответствует оценке, полученной при использовании $\hat{\lambda}$ в (29.29).

Обобщение модели структурной зависимости

29.13 Как мы уже отмечали после (29.18), модель структурной зависимости, рассмотренная в **29.9—12**, обладает тем ограничением, что все x_i имеют одно и то же среднее (откуда следует также равенство средних для y_i). Предполагалось, что выполнены условия

$$\mathbf{M}\xi_i = \mathbf{M}x_i = \mu \quad \text{для всех } i, \tag{29.33}$$

$$\mathbf{M}\eta_i = \mathbf{M}y_i = \alpha_0 + \alpha_1\mu \quad \text{для всех } i. \tag{29.34}$$

Если мы теперь ослабим условие (29.33) и предположим, что

$$\mathbf{M}\xi_i = \mathbf{M}x_i = \mu_i, \quad i = 1, 2, \ldots, n, \tag{29.35}$$

то (29.34) примет вид

$$\mathbf{M}\eta_i = \alpha_0 + \alpha_1\mu_i. \tag{29.36}$$

Это уже более содержательная модель, из которой можно получить функциональную, если положить $\sigma_x^2 = \sigma_y^2 = 0$, так что $X_i = \mu_i$, $Y_i = \alpha_0 + \alpha_1 X_i$.

Однако, принимая эту более общую модель, мы приходим к совсем другой задаче оценивания. Действительно, здесь все μ_i неизвестны, и поэтому вместо шести оцениваемых параметров, как в (29.18), мы имеем $n+5$ параметров. Существенно новым является то, что вместе с каждым новым наблюдением добавляется новый неизвестный параметр, так что неудивительно, что в данном случае мы сталкиваемся с новыми проблемами. Параметры, связанные с отдельными наблюдениями, были названы Нейманом и Скотт (1948) «несущественными» (incidental), а остальные параметры, общие для всех наблюдений, — «структурными» (structural). В примере 18.16 мы уже встречались с задачей, в которой фигурировали несущественные параметры.

Рассмотрим теперь процесс оценивания по методу МП при наличии несущественных параметров, причем начнем сразу со случая функциональной зависимости, который нас интересует больше всего.

Оценивание параметров функционального соотношения по методу МП

29.14 Предположим, что (29.4), (29.5) и (29.6) выполнены и что δ_i и ε_i — независимые нормально распределенные случайные величины. Поскольку X_i — математическая, а не случайная величина, то $\sigma_x^2 = 0$ и имеется $n+4$ параметров, а именно α_0, α_1, σ_δ^2, σ_ε^2, и n значений X_i. Функция правдоподобия в данном случае имеет вид

$$L \propto \sigma_\delta^{-n} \sigma_\varepsilon^{-n} \exp\left[-\frac{1}{2\sigma_\delta^2}\sum_i (\xi_i - X_i)^2 - \frac{1}{2\sigma_\varepsilon^2}\sum_i \{\eta_i - (\alpha_0 + \alpha_1 X_i)\}^2\right].$$

Дифференцируя $\log L$ по каждому из X_i и по остальным четырем параметрам, находим

$$\frac{\partial \log L}{\partial X_i} = \frac{\xi_i - X_i}{\sigma_\delta^2} + \frac{\alpha_1}{\sigma_\varepsilon^2}\{\eta_i - (\alpha_0 + \alpha_1 X_i)\} = 0, \quad i = 1, 2, \ldots, n, \tag{29.37}$$

$$\frac{\partial \log L}{\partial \alpha_0} = \frac{1}{\sigma_\varepsilon^2}\sum_i \{\eta_i - (\alpha_0 + \alpha_1 X_i)\} = 0, \tag{29.38}$$

$$\frac{\partial \log L}{\partial \alpha_1} = \frac{1}{\sigma_\varepsilon^2}\sum_i X_i \{\eta_i - (\alpha_0 + \alpha_1 X_i)\} = 0, \tag{29.39}$$

$$\frac{\partial \log L}{\partial \sigma_\delta} = -\frac{n}{\sigma_\delta} + \frac{1}{\sigma_\delta^3}\sum_i (\xi_i - X_i)^2 = 0, \tag{29.40}$$

$$\frac{\partial \log L}{\partial \sigma_\varepsilon} = -\frac{n}{\sigma_\varepsilon} + \frac{1}{\sigma_\varepsilon^3}\sum_i \{\eta_i - (\alpha_0 + \alpha_1 X_i)\}^2 = 0. \tag{29.41}$$

Суммируя (29.37) по i и используя (29.38), получаем

$$\sum_i (\xi_i - X_i) = 0.$$

Таким образом, если измерять ξ_i относительно их наблюденных средних, то мы получим МП-оценку для суммы X_i:

$$\left(\widehat{\sum_i X_i}\right) = \sum_i \xi_i = 0. \tag{29.42}$$

Из (29.38), учитывая (29.42), находим $\sum_i \eta_i = n\alpha_0$, или, если η_i измеряются относительно их наблюденных средних,

$$\hat{\alpha}_0 = 0. \tag{29.43}$$

Пользуясь (29.43), из (29.39) находим

$$\hat{\alpha}_1 = \sum \hat{X}_i \eta_i \Big/ \sum_i \hat{X}_i^2. \tag{29.44}$$

Уравнение (29.40) дает

$$\hat{\sigma}_\delta^2 = \frac{1}{n} \sum (\xi_i - \hat{X}_i)^2, \tag{29.45}$$

а (29.43) и (29.41) дают

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{n} \sum_i (\eta_i - \hat{\alpha}_i \hat{X}_i)^2. \tag{29.46}$$

Возводя (29.37) в квадрат и учитывая (29.43), получаем

$$\frac{(\xi_i - \hat{X}_i)^2}{\sigma_\delta^4} = \frac{\hat{\alpha}_1^2}{\sigma_\varepsilon^4} (\eta_i - \hat{\alpha}_1 \hat{X}_i)^2. \tag{29.47}$$

Суммируя затем (29.47) по i и используя отношение (29.45) и (29.46), мы приходим к заключению, что

$$\hat{\sigma}_\varepsilon^2 = \hat{\alpha}_1^2 \hat{\sigma}_\delta^2. \tag{29.48}$$

Подставляя (29.48) снова в (29.37) и исключая $\hat{\alpha}_1$, получаем

$$\frac{\xi_i - \hat{X}_i}{\hat{\sigma}_\delta^2} + \frac{\left(\eta_i - \frac{\hat{\sigma}_\varepsilon}{\hat{\sigma}_\delta} \hat{X}_i\right)}{\hat{\sigma}_\delta \hat{\sigma}_\varepsilon} = 0,$$

так что МП-оценка для X_i удовлетворяет уравнению

$$\hat{X}_i = \frac{1}{2}\left(\xi_i + \frac{\hat{\sigma}_\delta}{\hat{\sigma}_\varepsilon} \eta_i\right), \quad i = 1, 2, \ldots, n. \tag{29.49}$$

Чтобы получить МП-оценки для σ_δ^2 и σ_ε^2, надо решить $n+2$ уравнений (29.45), (29.46) и (29.49) относительно $n+2$ неизвестных X_i, $\hat{\sigma}_\delta^2$ и $\hat{\sigma}_\varepsilon^2$. Затем $\hat{\alpha}_1$ мы находим из (29.48).

Однако, как следует из соотношения (29.48), впервые полученного Линдли (1947), метод МП не годится для оценивания в рассматриваемом случае. Хотя мы априори ничего не знаем о значениях параметров α_1, σ_δ^2, σ_ε^2, (29.48) устанавливает определенное соотношение между их МП-оценками, которому сами параметры в рассматриваемой модели не обязаны удовлетворять. Фактически из (29.48) следует, что мы не можем оценить состоятельно все три параметра α_1, σ_δ^2, σ_ε^2. Таким образом, в данной ситуации метод МП неприемлем.

29.15 Как показали Нейман и Скотт (1948), в общем случае МП-оценки для структурных параметров не обязательно состоятельны, если в задаче имеются несущественные параметры. Позднее Кифером и Волфовицем (1956) было доказано, что *если несущественные параметры являются независимыми и одинаково распределенными случайными величинами*, а структурные параметры идентифицируемы, то в условиях регулярности МП-оценки структурных параметров состоятельны. Очевидно, выделенное курсивом условие возвращает нас от функциональной модели к структурной, которая рассматривалась в **29.9—12** и для которой при различных предположениях были получены МП-оценки для α_1. Еще раньше Нейман (1951) доказал существование состоятельных оценок для α_1 в модели структурной зависимости.

29.16 Из **29.14** следует, что без дополнительных предположений нельзя получить приемлемой МП-оценки для α_1 в случае функциональной зависимости. В действительности так же обстояло дело и в случае структурной зависимости, который в свете наших результатов и результатов, упомянутых в **29.15**, представляется существенно более простым. Эта необходимость в дополнительных предположениях часто удивляет исследователей, применяющих статистические методы. Обычно они переоценивают возможности получения с помощью статистических методов простых и приемлемых решений для любой просто формулируемой задачи. В связи с этим полезно прибегнуть к геометрической иллюстрации.

Рассмотрим точки (ξ_i, η_i), расположенные так, как изображено на рис. 29.1.

Любой наблюденной точке (ξ_i, η_i) соответствует «истинная» точка $(X_i, Y_i) = (\xi_i - \delta_i,\ \eta_i - \varepsilon_i)$, положение которой неизвестно. Поскольку в нашей модели δ_i и ε_i независимы и нормально распределены, то положение точки (ξ_i, η_i) равновероятно на любом эллипсе с центром в (X_i, Y_i) и осями, параллельными осям координат. И наоборот, если плотность распределения (ξ_i, η_i) симметрична по (ξ_i, η_i) и (X_i, Y_i), то существует эллиптическая доверительная область для (X_i, Y_i) с центром в (ξ_i, η_i) при любом заданном коэффициенте доверия. Эти области изо-

бражены на рис. 29.1. Эвристически задачу оценивания α_i можно понимать как задачу нахождения прямой, пересекающей возможно большее число таких доверительных областей. Трудность теперь ясна: неизвестны длины осей эллипсов — они зависят от параметров масштаба σ_δ, σ_ε. Понятно, что для того, чтобы сделать задачу определенной, достаточно знать эксцентриситет эллипсов, т. е. отношение $\sigma_\varepsilon/\sigma_\delta$. Напомним, что в случае структурной зависимости, рассмотренном в **29.9—10**, знания



Рис. 29.1. Доверительные области для (X_i, Y_i) (см. текст).

этого отношения оказалось достаточно для решения задачи оценивания.

29.17 Пусть, следовательно, отношение $\sigma_\varepsilon^2/\sigma_\delta^2 = \lambda$ известно. Заменяя σ_δ^2 на $\sigma_\varepsilon^2/\lambda$ в методе МП в **29.14**, мы видим, что несостоятельность, возникшая вследствие (29.48), исчезает, так как теперь требуется оценить только одну дисперсию ошибки, скажем σ_ε^2. Уравнения (29.40) и (29.41), которые приводили к (29.48), заменяются одним уравнением

$$\frac{\partial \log L}{\partial \sigma_\varepsilon} = \frac{-2n}{\sigma_\varepsilon} + \frac{\lambda}{\sigma_\varepsilon^3}\sum_i (\xi_i - X_i)^2 + \frac{1}{\sigma_\varepsilon^3}\sum \{\eta_i - (\alpha_0 + \alpha_1 X_i)\}^2 = 0,$$

из которого, учитывая остающиеся в силе соотношения (29.43) и (29.44), получаем

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{2n}\left\{\lambda \sum (\xi_i - X_i)^2 + \sum (\eta_i - \alpha_1 X_i)^2\right\}. \qquad (29.50)$$

Вместо (29.49) непосредственно из (29.37) находим

$$\lambda(\xi_i - \hat{X}_i) + \hat{\alpha}_1(\eta_i - \hat{\alpha}_1 \hat{X}_i) = 0,$$

или

$$\hat{X}_i = \frac{\lambda \xi_i + \hat{\alpha}_1 \eta_i}{\lambda + \hat{\alpha}_1^2}. \qquad (29.51)$$

Подставляя (29.51) в (29.44), получаем

$$\hat{\alpha}_1 = \frac{(\lambda + \hat{\alpha}_1^2)\{\lambda \sum \xi_i\eta_i + \hat{\alpha}_1 \sum \eta_i^2\}}{\lambda^2 \sum_i \xi_i^2 + \hat{\alpha}_1^2 \sum \eta_i^2 + 2\lambda\hat{\alpha}_1 \sum \xi_i\eta_i},$$

или, после упрощений,

$$\hat{\alpha}_1^2 \sum_i \xi_i\eta_i + \hat{\alpha}_1\left(\lambda \sum \xi_i^2 - \sum \eta_i^2\right) - \lambda \sum_i \xi_i\eta_i = 0. \qquad (29.52)$$

(29.52) отличается от (29.27) только незначительными изменениями обозначений. Таким образом, результат случая 3 из **29.12** сохраняется: если отношение дисперсий ошибок λ известно, то (29.29) дает МП-оценку для α_1 как в случае простейшего структурного соотношения, так и в случае линейного функционального соотношения.

29.18 Как было замечено в **29.10** (эти рассуждения применимы и в данном случае, поскольку мы оцениваем α_1 так же, как в случае 2 в **29.12**), оцененные прямые регрессии «ограничивают» оцененную структурную прямую. Это следует также из монотонной зависимости от λ оценки $\hat{\alpha}_1$, даваемой формулой (29.29) (в упражнении 29.1 читателю предлагается доказать этот факт). Таким образом, оцененные прямые регрессии дают математические границы для оцененной функциональной прямой. Однако эти границы могут быть слишком далеки друг от друга, чтобы иметь какое-нибудь практическое значение. В любом случае они, конечно, не являются ни в каком смысле вероятностными границами.

29.19 Знание отношения дисперсий ошибок позволило нам в **29.17** получить МП-оценки (29.29) и (29.50) для α_1 и σ_ε^2. Однако еще не все трудности преодолены. Дело в том, что если $\hat{\alpha}_1$ является состоятельной оценкой для α_1, то $\hat{\sigma}_\varepsilon^2$, как показал Линдли (1947), не будет таковой для σ_ε^2.

Чтобы доказать состоятельность α_1, заметим, что в соответствии с общими результатами главы 10 выборочные дисперсии и ковариация в (29.29) сходятся по вероятности к своим математическим ожиданиям. Таким образом, обозначая S_X^2 дисперсию ненаблюдаемой величины X_i, мы получаем (ср. (29.18) для случая структурной зависимости)

$$\left.\begin{aligned} s_\xi^2 &\to S_X^2 + \sigma_\delta^2 = S_X^2 + \frac{\sigma_\varepsilon^2}{\lambda}, \\ s_\eta^2 &\to \alpha_1^2 S_X^2 + \sigma_\varepsilon^2 = \alpha_1^2 S_X^2 + \lambda\sigma_\delta^2, \\ s_{\xi\eta} &\to \alpha_1 S_X^2. \end{aligned}\right\} \qquad (29.53)$$

Подставляя (29.53) в (29.29), получаем соотношение

$$\hat{\alpha}_1 \to \left(\{\alpha_1^2 S_X^2 + \lambda\sigma_\delta^2 - \lambda(S_X^2 + \sigma_\delta^2)\} + \right.$$
$$\left. + [\{\alpha_1^2 S_X^2 + \lambda\sigma_\delta^2 - \lambda(S_X^2 + \sigma_\delta^2)\}^2 + \lambda(\alpha_1 S_X^2)^2]^{1/2}\right) / \{2\alpha_1 S_X^2\} = \alpha_1, \quad (29.54)$$

устанавливающее состоятельность $\hat{\alpha}_1$. Приведенное доказательство проходит также в случае структурной зависимости, если всюду заменить S_X^2 на σ_x^2.

Несостоятельность $\hat{\sigma}_\varepsilon^2$ в случае функциональной зависимости доказывается так же просто. Подставляя (29.51) в (29.50), мы приходим к следующим выражениям для $\hat{\sigma}_\varepsilon^2$:

$$\hat{\sigma}_\varepsilon^2 = \frac{\lambda}{2(\lambda + \hat{\alpha}_1^2)} \cdot \frac{1}{n} \sum_i (\eta_i - \hat{\alpha}_1 \xi_i)^2, \quad (29.55)$$

$$\hat{\sigma}_\varepsilon^2 = \frac{\lambda}{2(\lambda + \hat{\alpha}_1^2)} (s_\eta^2 - 2\hat{\alpha}_1 s_{\xi\eta} + \hat{\alpha}_1^2 s_\xi^2). \quad (29.56)$$

Из (29.55), учитывая (29.53) и (29.54), получаем

$$\hat{\sigma}_\varepsilon^2 \to \frac{\lambda}{2(\lambda + \alpha_1^2)} \left\{ \alpha_1^2 S_X^2 + \sigma_\varepsilon^2 - 2\alpha_1^2 S_X^2 + \alpha_1^2 \left(S_X^2 + \frac{\sigma_\varepsilon^2}{\lambda} \right) \right\} = \frac{1}{2} \sigma_\varepsilon^2. \quad (29.57)$$

Доказанная несостоятельность МП-оценки напоминает несостоятельность, отмеченную в примере 18.16, которая была вызвана тем, что объем выборки был равен двум, а смещение МП-оценки имело порядок $1/n$. Здесь, по существу, σ_ε^2 также оценивается по парам (ξ_i, η_i), что легко видеть из (29.55). Следовательно, несостоятельность МП-оценки в данном случае вызвана влиянием смещенности МП-оценки при малом объеме выборки. Отмеченная несостоятельность, однако, не доставляет неприятностей: чтобы получить состоятельную оценку для σ_ε^2, надо заменить число наблюдений $2n$ числом степеней свободы $2n - (n+2) = n - 2$ в знаменателе $\hat{\sigma}_\varepsilon^2$. Состоятельная оценка, следовательно, равна $\frac{2n}{n-2} \hat{\sigma}_\varepsilon^2$.

Мы, таким образом, видим, что в случае функциональной зависимости для состоятельности МП-оценок всех структурных параметров недостаточно даже знания $\lambda = \sigma_\varepsilon^2 / \sigma_\delta^2$. Что касается структурных зависимостей, то в некоторых случаях состоятельность МП-оценок для структурных параметров гарантируется сформулированной в **29.15** теоремой Кифера — Волфовица.

Пример 29.1

Р. Браун (1957) приводит девять пар наблюдений

ξ:	1,8	4,1	5,8	7,5	9,3	10,6	13,4	14,7	18,9
η:	6,9	12,5	20,0	15,7	24,9	23,4	30,2	35,6	39,1

порожденных линейной функциональной зависимостью $Y=\alpha_0+\alpha_1 X$ с дисперсиями ошибок $\sigma_\delta^2=\sigma_\varepsilon^2$. Таким образом, мы имеем $\lambda=1$, $n=9$. Далее, вычисляем

$$\sum\xi=86{,}1, \quad \sum\eta=208{,}3,$$
$$\bar{\xi}=9{,}57, \quad \bar{\eta}=23{,}14$$

и (с точностью до трех значащих цифр)

$$ns_\xi^2=238, \quad ns_\eta^2=906, \quad ns_{\xi\eta}=451.$$

Используя затем (29.29), получаем

$$\hat{\alpha}_1=\frac{(906-238)+\{(906-238)^2+4\,(451)^2\}^{1/2}}{2\times 451}=\frac{668+1122}{902}=1{,}99.$$

Если принять за начало отсчета наблюденные средние, то в соответствии с (29.43) будем иметь $\hat{\alpha}_0=0$ и оцениваемая прямая примет вид

$$Y-23{,}14=1{,}99\,(X-9{,}57), \quad \text{или} \quad Y=1{,}99X+4{,}01.$$

Состоятельная оценка σ_ε^2 согласно **29.19** равна $s_\varepsilon^2=\frac{2n}{n-2}\,\hat{\sigma}_\varepsilon^2$, где $\hat{\sigma}_\varepsilon^2$ определяется формулой (29.56). Мы, следовательно, получаем

$$s_\varepsilon^2=\frac{9}{7}\,\frac{1}{1+\hat{\alpha}_1^2}\,(s_\eta^2-2\hat{\alpha}_1 s_{\xi\eta}+\hat{\alpha}_1^2 s_\xi^2)=$$
$$=\frac{1}{7\,(1+1{,}99^2)}\{906-(3{,}98\cdot 451)+(1{,}99)^2\cdot 238\}=1{,}53.$$

В действительности данные были порождены линейной функциональной зависимостью $(Y-\bar{\eta})=2\,(X-\bar{\xi})$ со случайными нормальными ошибками δ, ε, имеющими общую дисперсию $\sigma_\varepsilon^2=1$, так что найденные оценки, особенно $\hat{\alpha}$, оказались довольно хорошими даже при таком небольшом числе наблюдений, как 9.

Интервальное оценивание и критерии

29.20 До сих пор рассматривались только точечные оценки параметров. Теперь мы займемся интервальными оценками и в первую очередь найдем доверительные интервалы (и соответствующие критерии) только для одного α_1. Эта задача была

решена Кризи (1956) для случая, когда известно отношение дисперсий ошибок λ. Деля η на $\lambda^{1/2}$, мы приходим к случаю $\lambda = 1$, так что без потери общности можно считать дисперсии ошибок равными. В этом случае функция правдоподобия имеет вид

$$L \propto \sigma_\varepsilon^{-2n} \exp\left\{ -\frac{1}{2\sigma_\varepsilon^2}\left(\sum_{i=1}^{n}\delta_i^2 + \sum_{i=1}^{n}\varepsilon_i^2\right)\right\} \tag{29.58}$$

независимо от того, имеем мы дело со структурной или с функциональной зависимостью. Максимизация (29.58) равносильна минимизации выражения в круглых скобках, которое может быть записано как $\sum_{i=1}^{n}(\delta_i^2 + \varepsilon_i^2)$. По теореме Пифагора видим, что при оценивании по методу МП минимизируется сумма квадратов длин перпендикуляров, опущенных из наблюденных точек на оцененную прямую. Это ясно и интуитивно, если учесть равенство дисперсий ошибок.

Пользуясь обозначениями

$$\alpha_1 = \operatorname{tg}\theta, \qquad \hat{\alpha}_1 = \operatorname{tg}\hat{\theta}, \tag{29.59}$$

получаем из (29.29), учитывая инвариантность МП-оценок относительно преобразований, что МП-оценка для $\operatorname{tg} 2\theta$ равна

$$\operatorname{tg} 2\hat{\theta} = \frac{2\operatorname{tg}\hat{\theta}}{1-\operatorname{tg}^2\hat{\theta}} = \frac{2\hat{\alpha}_1}{1-\hat{\alpha}_1^2} = \frac{2s_{\xi\eta}}{\left|s_\xi^2 - s_\eta^2\right|}. \tag{29.60}$$

Благодаря использованию модуля в знаменателе правой части (29.60) знак $\operatorname{tg} 2\hat{\theta}$ совпадает со знаком $\hat{\alpha}_1$ и знаком $s_{\xi\eta}$.

В силу (29.18) ξ и η некоррелированы тогда и только тогда, когда $\alpha_1 = 0$, а поскольку ξ и η нормально распределены, то их выборочный коэффициент корреляции

$$r = s_{\xi\eta}/(s_\xi s_\eta)$$

в этом случае имеет распределение (16.62) или, что эквивалентно в силу (16.63), статистика

$$t = \{(n-2)\, r^2/(1-r^2)\}^{1/2} \tag{29.61}$$

имеет распределение Стьюдента с $n-2$ степенями свободы. Поскольку

$$\sin^2 2\hat{\theta} = \operatorname{tg}^2 2\hat{\theta}/(1+\operatorname{tg}^2 2\hat{\theta}),$$

то статистику (29.61) можно записать, учитывая (29.60), в виде

$$t = \left\{(n-2)\sin^2 2\hat{\theta}\left[\frac{(1/4)\left(s_\xi^2 - s_\eta^2\right) + s_{\xi\eta}^2}{s_\xi^2 s_\eta^2 - s_{\xi\eta}^2}\right]\right\}^{1/2}. \tag{29.62}$$

Эта статистика может быть использована для проверки гипотезы $\alpha_1 = \theta = 0$.

29.21 Если мы хотим проверить гипотезу о том, что α_1 имеет некоторое ненулевое значение, то возникает трудность, состоящая в том, что коэффициент корреляции между ξ и η, равный согласно (29.18) с $\sigma_\delta^2 = \sigma_\varepsilon^2$

$$\rho = \frac{\alpha_1 \sigma_x^2}{\left\{\left(\alpha_1^2 \sigma_x^2 + \sigma_\varepsilon^2\right)\left(\sigma_x^2 + \sigma_\varepsilon^2\right)\right\}^{1/2}},$$

зависит от неизвестной дисперсии σ_x^2 (под которой, как и до этого, в случае функциональной зависимости будет пониматься S_X^2). Чтобы избавиться от этой трудности, заметим, что при известном α_1 (и, следовательно, известном θ) можно перейти от наблюденных значений (ξ_i, η_i) к новым (ξ_i', η_i') с помощью преобразования

$$\left.\begin{aligned} \xi' &= \eta \sin\theta + \xi \cos\theta, \\ \eta' &= \eta \cos\theta - \xi \sin\theta, \end{aligned}\right\}$$

заключающегося в повороте осей координат на угол θ. Таким образом, вместо проверки гипотезы о том, что α_1 приняло некоторое значение, будет проверяться гипотеза $\alpha_1 = \theta = 0$ для преобразованных переменных. Поскольку дисперсии и ковариации инвариантны относительно ортогонального преобразования, то (29.62) будет по-прежнему нашей статистикой критерия, надо только заменить $\hat{\theta}$ на $(\hat{\theta} - \theta)$.

Остается еще та трудность, что каждому значению t в (29.62), измененному таким образом, соответствуют вследствие периодичности синуса четыре значения θ. Если можно предположить, что вероятность неравенства $|\hat{\theta} - \theta| > \pi/4$ пренебрежимо мала, то указанная трудность исчезает, и мы можем использовать (29.62) (с заменой θ на $(\hat{\theta} - \theta)$) как для проверки гипотезы о величине $\alpha_1 = \operatorname{tg}\theta$, так и для нахождения доверительных границ для θ и, следовательно, для α_1. Доверительные границы для θ имеют вид

$$\hat{\theta} \pm \frac{1}{2} \arcsin\left[2t\left\{\frac{s_\xi^2 s_\eta^2 - s_{\xi\eta}^2}{(n-2)\left[(s_\xi^2 - s_\eta^2)^2 + 4s_{\xi\eta}^2\right]}\right\}^{1/2}\right], \qquad (29.63)$$

где t — значение t-статистики Стьюдента, соответствующее $n-2$ степеням свободы и выбранному коэффициенту доверия. Поскольку было использовано условие $|\hat{\theta} - \theta| < \pi/4$, то данный метод годится только для больших выборок.

Пример 29.2

Для данных примера 29.1 имеем

$$\hat{\theta} = \operatorname{arctg} 1{,}99 = 0{,}35\pi.$$

Таблица распределения Стьюдента при числе степеней свободы 7 и коэффициенте доверия 0,95 дает $t = 2{,}36$. Следовательно, (29.63) с учетом результатов вычислений для $\hat{\alpha}_1$ принимает вид

$$0{,}35\pi \pm \frac{1}{2}\arcsin\left[4{,}72\,\frac{(238\times 906 - 451^2)^{1/2}}{\sqrt{7}\times 1122}\right] =$$

$$= 0{,}35\pi \pm \frac{1}{2}\arcsin 0{,}1742 = 0{,}35\pi \pm 0{,}03\pi.$$

Таким образом, 95-процентные доверительные границы для θ равны $0{,}32\pi$ и $0{,}38\pi$. Доверительные границы для α_1 равны просто тангенсам этих углов 1,58 и 2,53. Вследствие использования преобразования от α_1 к θ при нахождении этих границ МП-оценка 1,99 не находится точно в середине между этими границами. Поскольку число степеней свободы было небольшим, полученные пределы оказались довольно широкими.

29.22 Тем же способом, каким мы нашли доверительные границы для α_1 в **29.21**, мы можем найти доверительную область для всей прямой, *если известны обе дисперсии ошибок* (Р. Браун (1957)). Действительно, несмотря на доказанную в (29.9) коррелированность $(\varepsilon - \alpha_1\delta)$ с ξ, можно записать (29.8) в виде

$$\eta - (\alpha_0 + \alpha_1\xi) = \varepsilon - \alpha_1\delta. \tag{29.64}$$

Правая часть (29.64) распределена нормально со средним нуль и дисперсией $\sigma_\varepsilon^2 + \alpha_1^2\sigma_\delta^2$, а левая часть (29.64) содержит только наблюдаемые величины ξ, η и параметры α_0, α_1. Отсюда следует, что величина

$$\frac{\sum_{i=1}^{n}\{\eta_i - (\alpha_0 + \alpha_1\xi_i)\}^2}{\sigma_\varepsilon^2 + \alpha_1^2\sigma_\delta^2} \tag{29.65}$$

имеет распределение χ^2 с n степенями свободы. Если σ_δ^2 и σ_ε^2 известны, то мы без потери общности можем считать их равными единице, поскольку для этого достаточно разделить ξ_i на σ_δ и η_i на σ_ε. Тогда (29.65) примет вид

$$\chi_n^2 = \sum_{i=1}^{n}\frac{\{\eta_i - (\alpha_0 + \alpha_1\xi_i)\}^2}{1 + \alpha_1^2}. \tag{29.66}$$

Можно использовать (29.66) для нахождения доверительной области для прямой. Для этого из таблиц χ^2 определим c_γ такое, что

$$\mathsf{P}\{\chi_n^2 > c_\gamma\} = 1 - \gamma.$$

Тогда вероятность выполнения неравенства

$$\sum_{i=1}^{n} \{\eta_i - (\alpha_0 + \alpha_1 \xi_i)\}^2 \leqslant c_\gamma (1 + \alpha_1^2) \tag{29.67}$$

будет равна γ. Принимая, как и раньше, за начало отсчета наблюденные средние $\bar{\xi}$, $\bar{\mu}$, можно переписать (29.67) в виде

$$s_\eta^2 + \alpha_0^2 + \alpha_1^2 s_\xi^2 - 2\alpha_1 s_{\xi\eta} \leqslant c_\gamma (1 + \alpha_1^2)$$

или

$$\alpha_1^2 (s_\xi^2 - c_\gamma) - 2\alpha_1 s_{\xi\eta} + \alpha_0^2 \leqslant c_\gamma - s_\eta^2. \tag{29.68}$$

Соотношение (29.68) определяет область на плоскости (α_0, α_1), ограниченную кривой второго порядка. Эта область будет $100(1-\gamma)$-процентной доверительной областью для (α_0, α_1). Если увеличить c_γ (т. е. увеличить γ), то вновь полученная область будет содержать прежнюю.

Доверительная область будет ограниченной, если кривая второго порядка есть эллипс, и неограниченной, если она является гиперболой. При этом может оказаться, что вообще не существует действительных (α_0, α_1), удовлетворяющих (29.68). Эта трудность уже обсуждалась в другом контексте в примере 20.5.

Мы можем теперь точно так же, как это делалось в **28.26**, считать (29.68) ограничением, которому должна удовлетворять истинная прямая, и найти затем путем дифференцирования огибающую семейства всех возможных прямых, которая будет снова кривой второго порядка. В результате, как показал Р. Браун (1957), получится область на плоскости (X, Y), ограниченная кривой

$$\frac{(Y - \hat{\alpha}_1 X)^2}{c_\gamma - b_1} - \frac{(\hat{\alpha}_1 Y + X)^2}{b_2 - c_\gamma} = 1 + \hat{\alpha}_1^2, \tag{29.69}$$

где $\hat{\alpha}_1$ определяется формулой (29.29) и

$$b_1 = s_\xi^2 - \frac{s_{\xi\eta}}{\hat{\alpha}_1}, \quad b_2 = s_\xi^2 + \hat{\alpha}_1 s_{\xi\eta}.$$

$\hat{\alpha}_1$ предполагается положительным, так что $b_2 > b_1$. (29.69) дает требуемую доверительную область, которая будет гиперболой, если $b_1 < c_\gamma < b_2$, и эллипсом, если $c_\gamma > b_2$. Если $c_\gamma < b_1$, то мы не получим действительной кривой второго порядка.

29.23 В (29.69) использована оценка (29.29), которая может не быть МП-оценкой, если обе дисперсии ошибок известны, как это предполагалось в **29.22** (см. случай 4 в **29.12**).

Следует также обратить внимание на трудность другого рода. Нахождение доверительных границ для среднего значения нормального распределения с известной дисперсией с помощью распределения χ_n^2 не является эффективной процедурой. Это было отмечено в примере 20.5, а также в примере 25.3, где было показано, что АОЭ соответствующего критерия равна нулю. Поэтому не следует ожидать, что доверительная область (29.69), которая существенно основана на этой неэффективной процедуре, будет эффективной. Здесь она приведена лишь из-за отсутствия лучшей.

Линейная функциональная зависимость между несколькими переменными

29.24 Рассмотрим теперь задачу оценивания линейной функциональной зависимости k переменных. Чтобы сделать обозначения симметричными, мы будем использовать переменные X_1, X_2, ..., X_k и фиктивную переменную $X_0(\equiv 1)$, которые будут связаны соотношением

$$\sum_{j=0}^{k} \alpha_j X_j = 0. \tag{29.70}$$

Все переменные, кроме X_0, подвержены ошибкам, так что наблюдаемые величины ξ_j имеют вид

$$\xi_{ji} = X_{ji} + \delta_{ji}, \quad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, k. \tag{29.71}$$

Конечно, $\xi_{0i} = X_{0i} = 1$ для всех i. Как и раньше, мы предположим, что δ_{ji} нормально распределены, не зависят от X и друг от друга и имеют нулевые средние; кроме того, чтобы сделать параметры идентифицируемыми, будем предполагать, что отношения дисперсий ошибок известны. Если предположить, что

$$\frac{\mathsf{D}\delta_1}{\lambda_1} = \frac{\mathsf{D}\delta_2}{\lambda_2} = \ldots = \frac{\mathsf{D}\delta_k}{\lambda_k}, \tag{29.72}$$

где λ известны, то, деля ξ_j на $\sqrt{\lambda_j}$, можно вообще исключить λ_j из рассмотрения. Полученные нормированные переменные будут иметь одну и ту же неизвестную дисперсию, скажем σ_δ^2.

Логарифм функции правдоподобия случайных ошибок будет иметь тогда вид

$$\log L = \text{const} - nk \log \sigma_\delta - \frac{1}{2\sigma_\delta^2} \sum_{j=1}^{k} \sum_{i=1}^{n} (\xi_{ji} - X_{ji})^2. \tag{29.73}$$

Если рассматривать наши данные как n точек в k-мерном пространстве, то задача будет заключаться в определении гиперплоскости (29.70). Максимизация функции правдоподобия эквивалентна минимизации двойной суммы в (29.73), представляющей собой сумму квадратов расстояний от наблюденных точек ξ до оцененных точек X, как в (29.58). Эта сумма минимальна, когда оцененные точки X являются основаниями перпендикуляров, опущенных из точек ξ на гиперплоскость. Таким образом, МП-оценка гиперплоскости определяется так, чтобы сумма квадратов расстояний от наблюденных точек до этой плоскости была минимальна.

29.25 Сформулированная выше задача довольно часто встречается в математике. Так как расстояние точки $(\xi_{1i}, \xi_{2i}, \ldots, \xi_{ki})$ от гиперплоскости (29.70) равно

$$\sum_{j=0}^{k} \alpha_j \xi_{ji} \Big/ \left(\sum_{j=0}^{k} \alpha_j^2 \right)^{1/2},$$

то она заключается в минимизации суммы

$$S = \sum_{i=1}^{n} \left(\sum_{j=0}^{k} \alpha_j \xi_{ji} \right)^2 \Big/ \sum_{j=0}^{k} \alpha_j^2.$$

Удобно сформулировать эту задачу также следующим образом: минимизировать сумму

$$S' = \sum_i \left(\sum_j \alpha_j \xi_{ji} \right)^2 \tag{29.74}$$

при ограничении

$$\sum \alpha^2 = \text{const}.$$

В соответствии с (29.70) можно без потери общности считать константу равной 1. Тогда задача будет заключаться в нахождении безусловного минимума выражения

$$\sum_i \left(\sum_j \alpha_j \xi_{ji} \right)^2 - \mu \sum_j \alpha_j^2,$$

где μ — неопределенный множитель Лагранжа. Дифференцируя по α_l, получаем

$$\sum_i \xi_{li} \left(\sum_j \alpha_j \xi_{ji} \right) = \mu \alpha_l, \qquad l = 0, 1, \ldots, k. \tag{29.75}$$

Первое из этих уравнений, соответствующее $l = 0$ $(\xi_{0i} = 1)$, может быть отброшено, если совместить начало отсчета со средним величин ξ, т. е. выбрать его так, чтобы выполнялись соотношения

$$\sum_i \xi_{ji} = 0.$$

Обозначая c_{ij} выборочную ковариацию i-й и j-й вариант, из (29.75) получаем

$$\sum_{j=1}^{k} c_{lj}\alpha_j = \frac{\mu}{n}\alpha_l, \qquad l = 1, 2, \ldots, k. \tag{29.76}$$

Перенося все члены в левую часть и исключая α_j из этой системы уравнений, находим

$$\begin{vmatrix} c_{11} - \frac{\mu}{n} & c_{12} & c_{13} & \ldots & c_{1k} \\ c_{12} & c_{22} - \frac{\mu}{n} & c_{23} & \ldots & c_{2k} \\ c_{13} & c_{23} & c_{33} - \frac{\mu}{n} & \ldots & c_{3k} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ c_{1k} & c_{2k} & c_{3k} & \ldots & c_{kk} - \frac{\mu}{n} \end{vmatrix} = 0. \tag{29.77}$$

Если разделить i-ю строку на S_i, выборочное стандартное отклонение ξ_i, а j-й столбец — на S_j, и обозначить r_{ij} выборочные коэффициенты корреляции, то (29.77) примет вид

$$\begin{vmatrix} 1-\theta_1 & r_{12} & r_{13} & \ldots & r_{1k} \\ r_{12} & 1-\theta_2 & r_{23} & \ldots & r_{2k} \\ r_{13} & r_{23} & 1-\theta_3 & \ldots & r_{3k} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{1k} & r_{2k} & r_{3k} & \ldots & 1-\theta_k \end{vmatrix} = 0, \tag{29.78}$$

где $\theta_j = \mu/(nS_j^2)$.

Можно решить (29.78) относительно μ и найти затем α_j с помощью (29.76). На практике обычно используется итерационный процесс, при котором α_j вычисляются одновременно. Полученные решения будут МП-оценками для α_j.

Заметим, что (29.78) является уравнением k-й степени относительно μ и имеет k корней (которые неотрицательны, так как корреляционная матрица всегда неотрицательно определена). Нам нужен наименьший из этих корней. Действительно, если мы умножим l-е уравнение в (29.75) на α_l и сложим последние k уравнений, то найдем, что сумма в левой части будет равна (29.74). Следовательно, сумма

$$S' = \mu \sum_{l=1}^{k} \alpha_l^2 = \mu \tag{29.79}$$

должна быть минимальной.

29.26 Метод, предложенный в **29.22** для случая двух переменных, можно обобщить с целью получения поверхностей второго порядка, ограничивающих доверительные области для

$\alpha_1, \alpha_2, \ldots, \alpha_k$ в k-мерном пространстве. Аналогичным образом может быть найдена также поверхность второго порядка, «внутри» которой лежит гиперплоскость (29.70), представляющая функциональное соотношение. Конечно, такие области трудно представить наглядно и невозможно изобразить при $k > 3$. Этот вопрос подробно рассмотрен Р. Брауном и Фередеем (1958). (Некоторые из их результатов приведены в упражнениях 29.6—29.8.) Замечания, сделанные в **29.23**, остаются справедливыми и в этом случае.

Виллегас (1961) рассматривает случай, когда дисперсии ошибок неизвестны и оцениваются из повторных наблюдений (см. **29.12**) методом МП, и использует полученные результаты для нахождения доверительной области для линейного соотношения (1964). Его исследование охватывает также случай коррелированных ошибок.

Спрент (1966) дает общий метод оценивания коэффициентов, когда ошибки коррелированы.

29.27 До сих пор мы в основном имели дело с ситуациями, когда идентифицируемость обеспечивалась либо предположениями, касающимися дисперсий ошибок, либо повторными наблюдениями. Возникает вопрос, существует ли какой-нибудь другой путь, позволяющий добиться успеха в задаче оценивания линейной функциональной или структурной зависимости. Чтобы ответить на него, мы рассмотрим другие подходы к этой проблеме.

Метод смешанных семиинвариантов Гири

29.28 Рассмотрим сначала метод, предложенный Гири (1942b, 1943) для модели структурной зависимости, но который применим также и в случае функциональной зависимости.

Запишем линейное структурное соотношение в симметричной форме:

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k = 0. \tag{29.80}$$

Величины x_j наблюдаются с ошибками δ_j, так что в действительности мы наблюдаем величины $\xi_j = x_j + \delta_j$. Предполагается, что ошибки δ_j не зависят от x_j и друг от друга. Рассмотрим совместную производящую функцию семиинвариантов случайных величин ξ_j. Она равна сумме п. ф. с. x_j и п. ф. с. ошибок δ_j. Смешанные семиинварианты последних, согласно примеру 12.7, равны нулю. Следовательно, смешанные семиинварианты двух систем случайных величин ξ_j и x_j совпадают. Обозначая $\varkappa_x$ семиинварианты x_j и $\varkappa_\xi$ семиинварианты ξ_j и записывая порядки семиинвариантов в скобках, получаем

$$\varkappa_x(p_1, p_2, \ldots, p_k) = \varkappa_\xi(p_1, p_2, \ldots, p_k) \tag{29.81}$$

при условии, что по крайней мере два из p_i больше нуля. Таким образом, смешанные семиинварианты x_j могут быть оценены путем оценивания соответствующих смешанных семиинвариантов ξ_j.

29.29 Совместная х. ф. случайных величин x_j, измеренных относительно их истинных средних, равна

$$\varphi(t_1, t_2, \ldots, t_k) = \mathbf{M}\left\{\exp\left(\sum_{j=1}^{k} \theta_j x_j\right)\right\}, \qquad (29.82)$$

где $\theta_j = it_j$. Дифференцируя (29.82) по каждому из θ_j и используя (29.80), получаем

$$\sum_{j=1}^{k} \alpha_j \frac{\partial \varphi}{\partial \theta_j} = \mathbf{M}\left\{\left(\sum_j \alpha_j x_j\right)\exp\left(\sum_j \theta_j x_j\right)\right\} = 0. \qquad (29.83)$$

Отсюда для п. ф. с. $\psi = \log \varphi$ находим

$$\sum_j \alpha_j \frac{\partial \psi}{\partial \theta_j} = \frac{1}{\varphi} \sum_j \alpha_j \frac{\partial \varphi}{\partial \theta_j} = 0. \qquad (29.84)$$

Поскольку по определению

$$\psi = \sum \varkappa(p_1, p_2, \ldots, p_k) \frac{\theta_1^{p_1}\theta_2^{p_2} \ldots \theta_k^{p_k}}{p_1!\, p_2! \ldots p_k!},$$

то в силу (29.84) для всех $p_i \geqslant 0$ имеем

$$\alpha_1 \varkappa(p_1 + 1, p_2, \ldots, p_k) + \alpha_2 \varkappa(p_1, p_2 + 1, \ldots, p_k) + \ldots$$
$$\ldots + \alpha_k \varkappa(p_1, p_2, \ldots, p_k + 1) = 0. \qquad (29.85)$$

Соотношения (29.85) в силу (29.81) выполняются также для смешанных моментов наблюдаемых величин ξ, если по крайней мере два аргумента каждого семиинварианта больше нуля, т. е. если два или более $p_i > 0$. Приведенные рассуждения сохраняют силу и в случае функциональной зависимости. Случайная величина x_j, над которой было произведено n наблюдений, тогда заменяется набором из n фиксированных значений $X_{j1}, \ldots, X_{jn}$. Если эти значения рассматривать как конечную совокупность, которая извлечена без остатка, то ход рассуждений не меняется.

29.30 К сожалению, метод оценивания α_j с помощью (29.85) (с заменой смешанных семиинвариантов их оценками) совершенно непригоден в случае, когда x_j имеют совместное нормальное распределение, т. е. в наиболее важном для практики случае. Действительно, суммарный порядок каждого семиинварианта в (29.85) равен $\sum_{i=1}^{k} p_i + 1 \geqslant 3$, так как два или более $p_i > 0$. Однако, как было показано в **15.3**, все семиинварианты порядка

$\geqslant 3$ в случае нормального распределения равны нулю. Таким образом, в данной ситуации уравнения (29.85) оказываются бесполезными. Это неудивительно, так как мы, имея здесь дело со случаем неидентифицируемости, рассмотренным ранее в **29.9**, не сделали никаких дополнительных предположений, чтобы обеспечить идентифицируемость.

Даже в случае распределений, отличных от нормального, остается неясным, какие из соотношений (29.85) выбрать для оценивания k коэффициентов α_j. Нам нужно только k уравнений, но мы можем выбрать их бесконечным числом способов (в предположении, что все семиинварианты существуют). Естественно использовать уравнения наиболее низкого порядка, беря самые малые p_i, так как тогда оценки смешанных семиинвариантов в (29.85) будут иметь меньшие выборочные флуктуации (см. **10.8**(д)). Однако при этом следует соблюдать осторожность, даже в простейшем двумерном случае, который мы сейчас рассмотрим подробнее.

29.31 Рассмотрим случай $k = 2$, характеризуемый соотношением (29.13), которое мы запишем в виде $\alpha_1 x - y = 0$ и которое представляет собой частный случай (29.80) при $x \equiv x_1$, $y \equiv x_2$, $\alpha_2 = -1$, $\alpha_0 = 0$ (за начало отсчета приняты средние x и y). (29.85) приводит к соотношениям

$$\alpha_1 \varkappa(p_1 + 1,\ p_2) - \varkappa(p_1,\ p_2 + 1) = 0$$

или, если $\varkappa(p_1 + 1, p_2) \neq 0$,

$$\alpha_1 = \frac{\varkappa(p_1,\ p_2 + 1)}{\varkappa(p_1 + 1,\ p_2)}. \tag{29.86}$$

Последнее соотношение выполняется при любых p_1, $p_2 > 0$ и, следовательно, как замечено в **29.30**, бесполезно в случае нормального распределения. Однако даже если (ξ, η) распределены не по нормальному закону, их маргинальные распределения могут оказаться симметричными, и тогда все нечетные смешанные моменты и, следовательно, смешанные семиинварианты будут равны нулю. Таким образом, даже при отсутствии нормальности мы должны позаботиться о том, чтобы сумма $(p_1 + p_2 + 1)$ была четной. Приведем соотношения наиболее низкого порядка, которые обычно оказываются полезными:

$$\left.\begin{aligned} p_1 = 1,\quad p_2 = 2:\quad \alpha_1 &= \varkappa_{13}/\varkappa_{22}, \\ p_1 = 2,\quad p_2 = 1:\quad \alpha_1 &= \varkappa_{22}/\varkappa_{31}. \end{aligned}\right\} \tag{29.87}$$

В этих соотношениях используются семиинварианты (ξ, η), которые нужно оценить по наблюдениям.

Остается нерешенным вопрос о том, какое из соотношений (29.87) или их комбинацию использовать при решении конкретной задачи. Мадански (1959) предлагает находить линейную

комбинацию, имеющую минимальную дисперсию. Однако для этого требуются чрезвычайно громоздкие выкладки, и решение не всегда возможно, если нет дополнительных предположений о распределении (ξ, η).

Даже если симметрия отсутствует, исходное распределение может оказаться таким, что смешанные семиинварианты в знаменателе (29.87) будут равны или близки к нулю. В этом случае оценка будет иметь сильные выборочные флуктуации.

Пример 29.3

Рассмотрим данные примера 29.1 с новой точки зрения. Имеем

$$s_{21} = \sum (\xi - \bar{\xi})^2 (\eta - \bar{\eta}) = 445{,}853 = n\mu_{21},$$
$$s_{12} = \sum (\xi - \bar{\xi}) (\eta - \bar{\eta})^2 = 542{,}877 = n\mu_{12},$$
$$s_{31} = \sum (\xi - \bar{\xi})^3 (\eta - \bar{\eta}) = 24\,635{,}041 = n\mu_{31},$$
$$s_{22} = \sum (\xi - \bar{\xi})^2 (\eta - \bar{\eta})^2 = 46\,677{,}679 = n\mu_{22},$$

где $n = 9$. Используя (3.81), находим оценки семиинвариантов:

$$\varkappa_{21} = \mu_{21} = 49{,}539,$$
$$\varkappa_{12} = \mu_{12} = 60{,}320,$$
$$\varkappa_{31} = \mu_{31} - 3\mu_{20}\mu_{11} = -1232{,}45,$$
$$\varkappa_{22} = \mu_{22} - \mu_{20}\mu_{02} - 2\mu_{11}^2 = -2493{,}613.$$

Подставляя эти оценки в уравнение (29.86), получаем оценку для α_1:

$$p_1 = 1, \quad p_2 = 1: \quad \frac{\varkappa_{12}}{\varkappa_{21}} = \frac{60{,}320}{49{,}539} = 1{,}22. \tag{29.88}$$

Второе уравнение (29.87) дает более близкое к истине значение:

$$p_1 = 2, \quad p_2 = 1: \quad \frac{\varkappa_{22}}{\varkappa_{31}} = \frac{-2493{,}613}{-1232{,}45} = 2{,}02. \tag{29.89}$$

Может показаться более предпочтительным вместо семиинвариантов в этих уравнениях использовать k-статистики. Поскольку мы имеем дело с центральными моментами, то в соответствии с **13.2** получаем

$$k_{21} = \frac{ns_{21}}{(n-1)(n-2)}, \qquad k_{12} = \frac{ns_{12}}{(n-1)(n-2)},$$
$$k_{31} = \frac{n(n+1)s_{31} - 3(n-1)s_{11}s_{20}}{(n-1)(n-2)(n-3)}$$
$$k_{22} = \frac{n(n+1)s_{22} - 2(n-1)s_{11}^2 - (n-1)s_{20}s_{02}}{(n-1)(n-2)(n-3)},$$

Отсюда видно, что использование k-статистик вместо выборочных семиинвариантов не меняет значения оценки (29.88). Что касается оценки (29.89), то она теперь равна

$$\frac{k_{22}}{k_{31}} = \frac{-2308{,}79}{-1057{,}19} = 2{,}18. \qquad (29.90)$$

Следует напомнить, что эти данные были в действительности порождены нормальными случайными величинами, так что неудивительно, что оценка (29.88) далека от истинного значения (МП-оценка в примере 29.1 была равна 1,99). Это полностью согласуется с замечаниями, сделанными в **29.30—31**, так как, по существу, мы оцениваем здесь отношение двух нулевых величин.

Отметим, что оценка (29.89) оказалась немного ближе к МП-оценке, чем более точная на первый взгляд оценка (29.90). «Уточнение» это, однако, только кажущееся, так как хотя k-статистики являются несмещенными оценками для семиинвариантов, в данном случае оценивается отношение семиинвариантов, так что смещены обе оценки. Оценка (29.89) вычисляется несколько проще.

Читатель может проверить, что если использовать первое из уравнений (29.87), то мы получим $\mu_{13} = 10\,003$, $\varkappa_{13} = -5131$, откуда найдем оценку $\varkappa_{13}/\varkappa_{22} = 2{,}06$, которая очень близка к (29.89).

В случае больших выборок из нормального распределения ни одна из рассмотренных оценок не будет надежной.

29.32 Резюмируя, можно сказать, что метод оценивания α_1 с помощью смешанных семиинвариантов, хотя и не требует дополнительных предположений, оказывается легко уязвимым, причем с несколько неожиданной стороны. Оценкой α_1 служит отношение семиинвариантов, и если семиинвариант в знаменателе равен нулю или близок к нему, то оценка может иметь сильные выборочные флуктуации, которые не только не исчезают при возрастании объема выборки, а, наоборот, могут увеличиться.

Использование дополнительной информации: инструментальные переменные

29.33 Предположим теперь, что вместе с ξ и η наблюдается случайная величина ζ, коррелированная с ненаблюдаемой истинной переменной x и некоррелированная с ошибками наблюдений. Знание ζ, очевидно, дает дополнительную информацию о значении x, которую можно попытаться использовать. Переменная ζ называется *инструментальной*, поскольку она играет только вспомогательную роль при оценивании зависимости между y

и x. Мы будем считать, что ξ, η и ζ измеряются относительно своих выборочных средних.

Рассмотрим следующую оценку для α_1:

$$a_1 = \sum_{i=1}^{n} \zeta_i \eta_i \Big/ \sum_{i=1}^{n} \zeta_i \xi_i, \tag{29.91}$$

которую можно записать также в виде

$$a_1 \sum_{i=1}^{n} \zeta_i \xi_i = \sum_{i=1}^{n} \zeta_i \eta_i$$

или

$$a_1 \sum_i \zeta_i (x_i + \delta_i) = \sum_i \zeta_i (\alpha_0 + \alpha_1 x_i + \varepsilon_i). \tag{29.92}$$

Все выборочные ковариации в (29.92) сходятся по вероятности к своим математическим ожиданиям, поэтому в силу некоррелированности ζ с δ и ε из (29.92) получаем

$$a_1 \operatorname{cov}(\zeta, x) \to \alpha_1 \operatorname{cov}(\zeta, x). \tag{29.93}$$

В том и только том случае, когда

$$\lim_{n\to\infty} \operatorname{cov}(\zeta, x) \neq 0, \tag{29.94}$$

из (29.93) следует, что

$$a_1 \to \alpha_1, \tag{29.95}$$

т. е. оценка a_1 состоятельна. Заметим, что мы не делали относительно переменной ζ никаких предположений, кроме ее коррелированности с x и некоррелированности с ошибками. В частности, она может быть дискретной случайной величиной. В упражнении 29.16 выясняется вопрос об эффективности a_1. См. также упражнения 29.15, 29.16.

29.34 Независимо от вида инструментальной переменной она не только позволяет нам получить состоятельную оценку (29.91) для α_1, но также, как показал Дёрбин (1954), позволяет найти доверительные границы для (α_0, α_1).

В силу (29.8) $\eta - \alpha_0 - \alpha_1\xi = \varepsilon - \alpha_1\delta$. Поскольку переменная ζ некоррелирована с δ и ε, она некоррелирована и с $(\eta - \alpha_0 - \alpha_1\xi)$. Отсюда следует (см. **26.23**(а)), что при заданных α_0 и α_1 выборочный коэффициент корреляции r между ζ и $(\eta - \alpha_0 - \alpha_1\xi)$ распределен так, что

$$t^2 = (n-2)\, r^2/(1-r^2) \tag{29.96}$$

имеет t^2-распределение Стьюдента с $n-2$ степенями свободы. Обозначая $t^2_{1-\gamma}$ значение t^2, удовлетворяющее условию

$$\mathsf{P}\{t^2 \leqslant t^2_{1-\gamma}\} = 1 - \gamma,$$

мы получаем (учитывая монотонную зависимость $r^2 = t^2/\{t^2 + (n-2)\}$ от t^2)

$$\mathsf{P}\left\{\frac{[\sum \zeta(\eta - \alpha_0 - \alpha_1\xi)]^2}{\sum \zeta^2 \sum (\eta - \alpha_0 - \alpha_1\xi)^2} \leqslant r^2_{1-\gamma}\right\} = 1 - \gamma,$$

или

$$\mathsf{P}\left\{\frac{(\sum \zeta\eta)^2 - 2\alpha_1 \sum \zeta\eta \sum \zeta\xi + \alpha_1^2 (\sum \zeta\xi)^2}{\sum \zeta^2 (\sum \eta^2 + n\alpha_0^2 - 2\alpha_1 \sum \eta\xi + \alpha_1^2 \sum \xi^2)} \leqslant r^2_{1-\gamma}\right\} = 1 - \gamma. \quad (29.97)$$

Таким образом, если не считать зависимости от наблюдаемых величин ζ, η, ξ, соотношение (29.97) зависит только от α_0 и α_1. Оно определяет доверительную область на плоскости (α_0, α_1) с коэффициентом доверия $1 - \gamma$, ограниченную кривой второго порядка. Если параметр α_0 известен, то (29.97) дает доверительный интервал для α_1. В этом случае t^2 имеет $n - 1$ степеней свободы, поскольку оценивается только один параметр. Позже мы увидим, что для некоторых инструментальных переменных специального вида можно получить доверительный интервал для α_1 даже при неизвестном α_0.

29.35 Обычно при использовании инструментальных переменных возникает практическая трудность нахождения такой случайной величины, относительно которой было бы известно, что она коррелирована с x и некоррелирована с δ и ε. Лишь в редких случаях мы знаем систему настолько хорошо, чтобы быть уверенными в правильности предположений такого рода. Однако если в качестве инструментальной переменной мы берем дискретную «группирующую» величину (т. е. мы классифицируем наблюдения в соответствии с принадлежностью к определенным группам и рассматриваем эту классификацию как дискретную случайную величину), то выполнение указанных условий становится более вероятным. Действительно, из физических соображений может быть известно, что наблюдения разделяются на несколько групп, связанных с истинными значениями x, тогда как ошибки наблюдений никак не связаны с классификацией. Рассмотрим, например, зависимость между давлением и объемом, приведенную в **29.1**, и предположим, что соотношение (29.2) выполнено. Логарифмируя, получаем

$$\log P = C - \gamma \log V.$$

Последнее соотношение принадлежит к рассматриваемому нами типу с $\alpha_0 = C$ и $\alpha_1 = \gamma$. (См., однако, **29.55**.)

Пусть известно, что определение объема проводилось то одним методом, то другим, причем метод 1 дает результаты, несколько отличающиеся от результатов, полученных с помощью метода 2. Классификация метод 1 — метод 2 будет тогда корре-

лирована с определением объема. Ошибки же определения объема и, конечно, давления (которое вообще определяется одним методом для всех наблюдений) могут быть совершенно некоррелированными с этой классификацией. Таким образом, мы получаем инструментальную переменную специального вида, связанную с группировкой на два класса.

Инструментальные переменные такого вида мы изучим подробнее.

Две группы равного объема

29.36 Предположим, что имеется четное число n наблюдений, которые разделены на две группы, состоящие из $n/2 = m$ наблюдений каждая. (Вопрос о том, как должно проводиться распределение по группам, будет рассмотрен чуть позже.) Пусть $\bar{\xi}$ — выборочное среднее ξ в первой группе, а $\bar{\xi}'$ — во второй группе. Аналогично определяются $\bar{\eta}$ и $\bar{\eta}'$. Мы можем оценить α_1, полагая инструментальную переменную ζ равной $+1$ для каждого наблюдения из первой группы и -1 для каждого наблюдения из второй группы. (29.91) в этом случае принимает вид

$$a_1 = \frac{\bar{\eta}' - \bar{\eta}}{\bar{\xi}' - \bar{\xi}}, \tag{29.98}$$

откуда в соответствии с принятой моделью получаем также оценку для α_0:

$$a_0 = (\bar{\eta}' + \bar{\eta}) - a_1(\bar{\xi}' + \bar{\xi}). \tag{29.99}$$

Геометрически процедура означает, что мы делим точки на плоскости (ξ, η) на две равные группы в соответствии со значением ξ и определяем центр тяжести каждой группы. Оценкой для истинной линейной зависимости служит тогда прямая, соединяющая эти центры тяжести.

Вальд (1940), получивший этот результат, показал, что оценка (29.98) для α_1 состоятельна, если группировка не зависит от ошибок и если значения x удовлетворяют соотношению

$$\lim_{n\to\infty} \inf |\bar{x}' - \bar{x}| > 0, \tag{29.100}$$

которое совпадает с (29.94), поскольку в данном случае $\operatorname{cov}(\xi, x) = \bar{x}' - \bar{x}$. Условие (29.100), очевидно, не будет выполнено, если наблюдения разделены на две группы случайным образом, так как тогда $\lim_{n\to\infty} |\bar{x}' - \bar{x}| = 0$. (См. упражнение 28.16, в котором рассматривается простая линейная модель.) Неудовлетворительной будет также классификация, относящая m наименьших значений ξ к одной группе, а m наибольших — к другой.

Нейман и Скотт (1951) показали, что в этом случае оценка не будет состоятельной. Действительно, легко видеть, что при такой классификации группировка в общем случае будет зависеть от ошибок. Из сказанного вытекает, что метод Вальда представляет интерес только тогда, когда имеется априорная информация (вроде упомянутой в **29.35**), обосновывающая (29.100).

29.37 Оценку (29.98) можно использовать для оценивания дисперсий ошибок. Действительно, поскольку в силу (29.18)

$$\sigma_\delta^2 = \mathsf{D}\xi - \frac{\operatorname{cov}(\xi, \eta)}{\alpha_1}, \quad \sigma_\varepsilon^2 = \mathsf{D}\eta - \alpha_1 \operatorname{cov}(\xi, \eta), \qquad (29.101)$$

то достаточно подставить состоятельные оценки s_ξ^2, s_η^2 и $s_{\xi\eta}$ вместо дисперсий и ковариаций (умножив их на $\frac{n}{n-1}$ для устранения смещения) и a_1 вместо α_1, чтобы получить нужные оценки

$$\left.\begin{aligned} s_\delta^2 &= \frac{n}{n-1}\left(s_\xi^2 - \frac{s_{\xi\eta}}{a_1}\right), \\ s_\varepsilon^2 &= \frac{n}{n-1}\left(s_\eta^2 - a_1 s_{\xi\eta}\right). \end{aligned}\right\} \qquad (29.102)$$

Пример 29.4

Применим в чисто иллюстративных целях этот метод к данным примера 29.1, где имеется девять наблюдений. Мы опустим медианное значение ξ и разделим оставшиеся наблюдения на две группы:

ξ:	1,8	4,1	5,8	7,5;	10,6	13,4	14,7	18,9
η:	6,9	12,5	20,0	15,7;	23,4	30,2	35,6	39,1

Вычисляем

$$\bar{\xi} = 19{,}2/4 = 4{,}800, \quad \bar{\xi}' = 57{,}6/4 = 14{,}400,$$
$$\bar{\eta} = 55{,}1/4 = 13{,}775, \quad \bar{\eta}' = 128{,}3/4 = 32{,}075,$$

откуда находим оценку

$$a_1 = \frac{32{,}075 - 13{,}775}{14{,}400 - 4{,}800} = 1{,}91,$$

достаточно близкую к истинному значению 2. Для взятых восьми наблюдений вычисляем $s_\xi^2 = 29{,}735$, $s_\eta^2 = 112{,}709$, $s_{\xi\eta} = 56{,}764$. Подставляя полученные значения в (29.102), получаем оценки

$$s_\delta^2 = -0{,}054, \quad s_\varepsilon^2 = 5{,}16,$$

которые оказываются очень плохими (истинные значения равны 1). Значение s_δ^2 даже отрицательно и, следовательно, «невозможно».

Легко видеть, что использованный метод действительно вполне может привести к неточным оценкам для дисперсий ошибок. Если истинные значения (x, y) отстоят друг от друга довольно далеко по сравнению с ошибками наблюдений, то наблюденные значения (ξ, η) будут сильно коррелированы и две линии регрессии будут близки одна к другой, а оценка a_1 будет близка как к коэффициенту регрессии η по ξ, равному $s_{\xi\eta}/s_\xi^2$, так и к величине $s_{\xi\eta}/s_\eta^2$, обратной коэффициенту регрессии ξ по η. Таким образом, в соответствии с (29.102) как s_δ^2, так и s_ε^2 будут близки к нулю и совсем незначительные изменения a_1 будут приводить к сильным изменениям этих оценок. В данном примере корреляция между ξ и η равна 0,98, и поэтому даже небольшого отклонения a_1 от истинного значения α_1 оказалось достаточно, чтобы сильно сдвинуть s_δ^2 вниз, а s_ε^2 вверх.

29.38 Вальд (1940) получил также доверительный интервал для α_1, применив следующую процедуру. Для каждой из двух групп вычисляются суммы квадратов и произведений относительно их средних и определяются объединенные оценки, каждая из которых основана на $(m-1)+(m-1)=n-2$ степенях свободы,

$$\left.\begin{aligned} S_\xi^2 &= \frac{1}{n-2}\left\{\sum_{i=1}^{m}(\xi_i-\bar\xi)^2+\sum_{i=1}^{m}(\xi_i'-\bar\xi')^2\right\},\\ S_\eta^2 &= \frac{1}{n-2}\left\{\sum_{i=1}^{m}(\eta_i-\bar\eta)^2+\sum_{i=1}^{m}(\eta_i'-\bar\eta')^2\right\},\\ S_{\xi\eta} &= \frac{1}{n-2}\left\{\sum_{i=1}^{m}(\xi_i-\bar\xi)(\eta_i-\bar\eta)+\sum_{i=1}^{m}(\xi_i'-\bar\xi')(\eta_i'-\bar\eta')\right\}. \end{aligned}\right\} \quad (29.103)$$

Эти три оценки в предположении нормальности распределены независимо от средних $\bar\xi$, $\bar\xi'$, $\bar\eta$, $\bar\eta'$ и, следовательно, независимо от оценки (29.98). Подставляя (29.103) в (29.101), мы получаем следующие случайные величины (еще зависящие от α_1):

$$S_\delta^2 = S_\xi^2 - S_{\xi\eta}/\alpha_1, \qquad S_\varepsilon^2 = S_\eta^2 - \alpha_1 S_{\xi\eta}. \quad (29.104)$$

Рассмотрим теперь сумму

$$\begin{aligned} S^2 = S_\varepsilon^2 + \alpha_1^2 S_\delta^2 &= S_\eta^2 + \alpha_1^2 S_\xi^2 - 2\alpha_1 S_{\xi\eta} = \\ &= \frac{1}{n-2}\left[\sum_{i=1}^{m}\{(\eta_i-\alpha_0-\alpha_1\xi_i)-(\bar\eta-\alpha_0-\alpha_1\bar\xi)\}^2+\right.\\ &\left.+\sum_{i=1}^{m}\{(\eta_i'-\alpha_0-\alpha_1\xi_i')-(\bar\eta'-\alpha_0-\alpha_1\bar\xi')\}^2\right]. \end{aligned} \quad (29.105)$$

Мы видим, что $(n-2)S^2$ является суммой двух сумм квадратов m центрированных независимых нормально распределенных случайных величин $(\eta_i - \alpha_0 - \alpha_1\xi_i)$. Из (29.8) следует, что каждая из них имеет дисперсию $\sigma_\varepsilon^2 + \alpha_1^2\sigma_\delta^2$. Таким образом, случайная величина

$$\frac{(n-2)\,S^2}{\sigma_\varepsilon^2 + \alpha_1^2\sigma_\delta^2}$$

имеет распределение χ^2 с $n-2$ степенями свободы. Определим теперь случайную величину

$$u = \frac{1}{2}(\bar{\xi}' - \bar{\xi})(a_1 - \alpha_1) = \frac{1}{2}\{(\bar{\eta}' - \bar{\eta}) - \alpha_1(\bar{\xi}' - \bar{\xi})\} =$$

$$= \frac{1}{2}\{(\bar{\eta}' - \alpha_0 - \alpha_1\bar{\xi}') - (\bar{\eta} - \alpha_0 - \alpha_1\bar{\xi})\} = \frac{1}{2}\{(\bar{\varepsilon}' - \alpha_1\bar{\delta}') - (\bar{\varepsilon} - \alpha_1\bar{\delta})\}.$$

Две компоненты в правой части, будучи функциями от средних значений ошибок в двух разных группах, распределены независимо. Таким образом, случайная величина u распределена нормально со средним нуль и дисперсией $\frac{1}{4}\cdot 2\,\frac{(\sigma_\varepsilon^2 + \alpha_1^2\sigma_\delta^2)}{m} = \frac{1}{n}(\sigma_\varepsilon^2 + \alpha_1^2\sigma_\delta^2)$. Кроме того, u зависит только от $\bar{\xi}'$, $\bar{\xi}$, $\bar{\eta}'$ и $\bar{\eta}$ и, следовательно, распределена независимо от S^2. Отсюда следует, что статистика

$$t = \frac{un^{1/2}}{S} = \frac{(\bar{\xi}' - \bar{\xi})(a_1 - \alpha_1)\,n^{1/2}}{2\left(S_\eta^2 - 2\alpha_1 S_{\xi\eta} + \alpha_1^2 S_\xi^2\right)^{1/2}} \tag{29.106}$$

имеет распределение Стьюдента с $n-2$ степенями свободы. Для любого заданного коэффициента доверия $1-\gamma$ мы имеем

$$\mathbf{P}\{t^2 \leqslant t_{1-\gamma}^2\} = 1 - \gamma. \tag{29.107}$$

Крайними значениями α_1, для которых выполняется условие (29.107), согласно (29.106) являются корни уравнения

$$(\bar{\xi}' - \bar{\xi})^2(a_1 - \alpha_1)^2 = \frac{4t_{1-\gamma}^2}{n}(S_\eta^2 - 2\alpha_1 S_{\xi\eta} + \alpha_1^2 S_\xi^2),$$

или

$$\alpha_1^2\left\{\frac{4t_{1-\gamma}^2}{n}S_\xi^2 - (\bar{\xi}' - \bar{\xi})^2\right\} + 2\alpha_1\left\{a_1(\bar{\xi}' - \bar{\xi})^2 - \frac{4t_{1-\gamma}^2}{n}S_{\xi\eta}\right\} +$$

$$+ \left\{\frac{4t_{1-\gamma}^2}{n}S_\eta^2 - a_1^2(\bar{\xi}' - \bar{\xi})^2\right\} = 0. \tag{29.108}$$

Это квадратное уравнение относительно α_1 с дискриминантом, равным

$$\left(\frac{4t_{1-\gamma}^2}{n}\right)^2(S_{\xi\eta}^2 - S_\xi^2 S_\eta^2) + \frac{4t_{1-\gamma}^2}{n}(a_1^2 S_\xi^2 - 2a_1 S_{\xi\eta} + S_\eta^2). \tag{29.109}$$

Первое слагаемое в (29.109) согласно неравенству Коши отрицательно, а второе положительно, так как множитель в скобках равен $\frac{1}{n-2}\sum_i (\eta_i - a_i\xi_i)^2$. Если n достаточно велико, то положительное слагаемое, имея в своем составе множитель $\frac{4t^2_{1-\gamma}}{n}$, будет больше отрицательного слагаемого, имеющего множитель $\left(\frac{4t^2_{1-\gamma}}{n}\right)^2$. В этом случае (29.108) будет иметь два действительных корня, являющихся доверительными границами для α_1.

29.39 Аналогично можно найти доверительную область для (α_0, α_1). Возьмем в качестве оценки для α_0 оценку a_0, определенную формулой (29.99), и рассмотрим случайную величину

$$v = a_0 - \alpha_0 = (\bar\eta' + \bar\eta) - \alpha_0 - \alpha_1(\bar\xi' + \bar\xi), \qquad (29.110)$$

которая нормально распределена с нулевым средним и дисперсией

$$\frac{1}{n}\mathsf{D}(\eta - \alpha_0 - \alpha_1\xi) = \frac{1}{n}(\sigma_\varepsilon^2 + \alpha_1^2\sigma_\delta^2),$$

равной дисперсии u в **29.38**. Легко показать, что v, как и u, распределена независимо от S^2, так что, заменяя u в (29.106) на v, мы получим t-статистику Стьюдента с $n-2$ степенями свободы. Если параметр α_1 известен, то можно использовать эту статистику для нахождения доверительного интервала для α_0. Это делается просто, так как α_0 входит только в числитель t. Однако этот случай не имеет большого практического значения, поскольку редко бывает, что мы знаем α_1 и не знаем α_0.

Можно, однако, показать, что u и v распределены независимо. Учитывая определения u и v, для этого достаточно показать, что

$$2u = (\bar\eta' - \bar\eta) - \alpha_1(\bar\xi' - \bar\xi)$$

не зависит от

$$\alpha_0 + v = (\bar\eta' + \bar\eta) - \alpha_1(\bar\xi' + \bar\xi).$$

Последние две случайные величины нормально распределены, причем первая из них имеет нулевое среднее. Их ковариация равна

$$\mathsf{M}(\bar\eta' + \bar\eta)(\bar\eta' - \bar\eta) + \alpha_1^2\mathsf{M}(\bar\xi' + \bar\xi)(\bar\xi' - \bar\xi) - 2\alpha_1\mathsf{M}(\bar\eta'\bar\xi' - \bar\eta\bar\xi).$$

Каждое из первых двух математических ожиданий есть математическое ожидание разности одинаково распределенных квадратов и, следовательно, равно нулю. Третье математическое ожидание есть математическое ожидание разности одинаково распределенных произведений и тоже равно нулю. Таким образом,

ковариация равна нулю и рассматриваемые случайные величины независимы, а следовательно, независимы также u и v.

Из доказанного следует, что статистика

$$\frac{u^2+v^2}{\left(\sigma_\varepsilon^2+\alpha_1^2\sigma_\delta^2\right)/n}$$

имеет распределение χ^2 с 2 степенями свободы, и поэтому статистика

$$F=\frac{(1/2)\,(u^2+v^2)}{S^2/n} \tag{29.111}$$

имеет распределение дисперсионного отношения с $(2, n-2)$ степенями свободы. Этот факт позволяет найти доверительную область для α_0 и α_1, которая (см. упражнение 29.5) представляет собой эллипс, как этого и следовало ожидать, принимая во внимание независимость и нормальность случайных величин u и v и их линейную зависимость от α_1 и α_0.

Полученная доверительная область не совпадает с доверительной областью, получаемой из (29.97) при $\zeta=\pm 1$. Только что рассмотренная доверительная область основана на распределении случайной величины $F=\frac{n}{2}\frac{(u^2+v^2)}{S^2}$, тогда как случайная величина в (29.97), числитель которой зависит только от u^2, не является монотонной функцией от F. Интуитивно кажется, что последняя область должна приводить к лучшему интервалу, однако нам не известно никаких результатов, относящихся к этому вопросу.

Три группы

29.40 Нэйр и Шривастава (1942), а также Бартлетт (1949) обратили внимание на то, что эффективность метода группировки может быть увеличена, если использовать вместо двух три группы и брать в качестве оценки α_1 угол наклона прямой, соединяющей центры тяжести двух крайних групп. (Мы уже фактически поступили так в примере 29.4, отбросив центральное наблюдение.) Метод трех групп может быть описан следующим образом.

Разделим n наблюдений на три группы, первая из которых будет содержать np_1 наблюдений, а третья — np_2. Вопрос о выборе значений p_1 и p_2 будет рассмотрен ниже. Метод двух групп представляет собой частный случай, соответствующий $p_1=p_2=1/2$ при четном n (средняя группа не содержит ни одной точки) и $p_1=p_2=\frac{n-1}{2n}$ при нечетном n (как было в примере 29.4). Рассматриваемая группировка соответствует инструментальной переменной ζ в (29.91), принимающей значения

$+1, 0$ и -1 для третьей, второй и первой групп соответственно. Оценка, как и раньше, имеет вид

$$a_1 = \frac{\bar{\eta}' - \bar{\eta}}{\bar{\xi}' - \bar{\xi}}, \qquad (29.112)$$

с той лишь разницей, символы со штрихами относятся к третьей группе, а без штрихов — к первой. Эта оценка будет состоятельной при тех же предположениях, что и раньше.

Нэйр и Шривастава (1942) и Бартлетт (1949) брали $p_1 = p_2 = 1/3$. В этом случае, аналогично **29.38**, мы определим S_ξ^2, S_η^2 и $S_{\xi\eta}$, объединяя внутригрупповые наблюденные дисперсии и ковариации и деля на число степеней свободы, равное теперь $n - 3$. Как и прежде, (29.104) определяет S_δ^2, и S_ε^2, и $S^2 = S_\varepsilon^2 + \alpha_1^2 S_\delta^2$. Статистика $(n-3)\,S^2/(\sigma_\varepsilon^2 + \alpha_1^2\sigma_\delta^2)$ имеет распределение χ^2 с $n-3$ степенями свободы, а статистика

$$(\bar{\xi}' - \bar{\xi})(a_1 - \alpha_1)\left(\frac{n}{6}\right)^{1/2}$$

не зависит от S^2 и имеет нормальное распределение с нулевым средним и дисперсией $\sigma_\varepsilon^2 + \alpha_1^2\sigma_\delta^2$. Таким образом, аналогом (29.106) будет статистика

$$t = \left(\frac{n}{6}\right)^{1/2} \frac{(\bar{\xi}' - \bar{\xi})(a_1 - \alpha_1)}{S}, \qquad (29.113)$$

имеющая распределение Стьюдента с $n-3$ степенями свободы, и мы сможем, как и раньше, найти доверительные интервалы, используя (29.113). Аналогично обобщаются на случай трех групп также результаты **29.39**.

29.41 Вопрос об оптимальном выборе p_1 и p_2 изучался для различных распределений x в предположении, что x измеряется *без ошибки*. В упражнении 29.11 приведен результат Бартлетта для прямоугольного распределения. Другие случаи рассматривались Тейлом и Ван Изереном (1956) и Гибсоном и Джоуэттом (1957). В целом они указывают на то, что для довольно широкого класса симметричных распределений x следует брать приблизительно $p_1 = p_2 = 1/3$, что обеспечивает эффективность 80—85% по сравнению с НК-оценкой, дающей минимальную дисперсию.

Данные об относительной эффективности двухгруппового и трехгруппового методов в случае наблюдений с ошибками неполны и противоречивы. Нэйр и Банерджи (1942) на основании выборочных экспериментов сделали вывод о том, что метод трех групп более эффективен. Однако пример, данный Мадански (1959), приводит к прямо противоположному заключению.

Пример 29.5

Примененный к данным примера 29.1 метод трех групп с $p_1 = p_2 = 1/3$ приводит к трем наблюдениям в каждой группе. Находим

$$\bar{\xi}' = 15{,}67, \quad \bar{\xi} = 3{,}90,$$
$$\bar{\eta}' = 34{,}97, \quad \bar{\eta} = 13{,}13,$$

откуда

$$a_1 = \frac{34{,}97 - 13{,}13}{15{,}67 - 3{,}90} = 1{,}86.$$

Полученное значение близко к значению 1,91, найденному в примере 29.4 методом двух групп, но отстоит несколько дальше от истинного значения 2.

Гальперин (1961b) дает обобщение (29.108) и (29.111), при котором требуется не группировка, а предположительное знание индивидуальных значений x_i, позволяющее максимизировать вероятность получения узких доверительных границ или областей. Этот метод особенно ценен при малых n.

Использование рангов

29.42 Завершая обсуждение методов группировки и инструментальных переменных, рассмотрим использование рангов. Предположим, что мы можем установить истинный порядок расположения отдельных наблюдений, исходя из порядка наблюденных значений одной из переменных. То есть мы предполагаем теперь, что не только две или три группы могут быть правильно упорядочены, но что значения x настолько удалены друг от друга по сравнению с их дисперсиями ошибок, что порядок наблюденных значений ξ совпадает с порядком ненаблюденных истинных значений x. Мы будем считать, что индексы соответствуют порядку наблюдений. Мы по-прежнему будем считать ошибки независимыми и предположим, что число наблюдений четно и равно $2m = n$. Каждой паре (ξ_i, η_i) соответствует пара (ξ_{m+i}, η_{m+i}). В качестве оценки для α_1 можно взять среднее или медиану следующих m оценок для α_1:

$$a(i) = \frac{\eta_{m+i} - \eta_i}{\xi_{m+i} - \xi_i}, \quad i = 1, 2, \ldots, m. \tag{29.114}$$

С другой стороны, можно рассмотреть все возможные оценки вида

$$a(i, j) = \frac{\eta_i - \eta_j}{\xi_i - \xi_j}, \quad i, j = 1, 2, \ldots, m. \tag{29.115}$$

Число таких оценок равно $\frac{1}{2}n(n-1)$. В качестве окончательной оценки для α_1 снова можно взять среднее или медиану всех таких оценок.

В этих методах, принадлежащих Тейлу (1950), очевидно, используется больше информации, чем в ранее рассмотренных методах группировки. Использование медианы вместо среднего позволяет, как будет показано в **29.43**, довольно легко строить доверительные интервалы.

Пример 29.6

Обращаясь снова к данным примера 29.1 и выбрасывая среднее значение, находим четыре оценки вида (29.114):

$$\frac{23{,}4-6{,}9}{10{,}6-1{,}8}=1{,}875, \quad \frac{30{,}2-12{,}5}{13{,}4-4{,}1}=1{,}903,$$

$$\frac{35{,}6-20{,}0}{14{,}7-5{,}8}=1{,}753, \quad \frac{39{,}1-15{,}7}{18{,}9-7{,}5}=1{,}887.$$

Медиана (полусумма средних значений) равна 1,88. Среднее равно 1,85.

В случае применения формулы (29.115) можно использовать все девять наблюдений. Мы получим следующие значения для $a(i, j)$: —2,529; —1,154; 0,708; 0,833; 0,941; 1,293; 1,342; 1,400; 1,458; 1,479; 1,544; 1,618; 1,677; 1,753; 1,797; 1,875; 1,883; 1,892; 1,903; 1,981; 2,009; 2,053; 2,179; 2,225; 2,385; 2,400; 2,429; 2,435; 2,458; 2,484; 2,764; 2,976; 3,275; 4,154; 4,412; 5,111. Медиана (среднее между восемнадцатым и девятнадцатым значениями) равна 1,90. Среднее равно 1,93.

29.43 Теперь мы откажемся от условия нормальности ошибок и наложим более слабое ограничение на разности $(\varepsilon_i - \alpha_1\delta_i)$, заключающееся в предположении, что они имеют одно и то же непрерывное распределение для всех i. Пользуясь терминологией **28.7**, можно сказать, что ошибки $\eta - \alpha_0 - \alpha_1\xi$ однородны (кроме того, добавляется непрерывность). Отсюда следует, что вероятность того, что одно значение, скажем $\varepsilon_i - \alpha_1\delta_i$, превосходит другое $\varepsilon_{m+i} - \alpha_1\delta_{m+i}$, равна 1/2.

Поскольку в силу (29.114)

$$a(i)=\frac{\eta_{m+i}-\eta_i}{\xi_{m+i}-\xi_i}=\alpha_1+\frac{(\eta_{m+i}-\alpha_1\xi_{m+i})-(\eta_i-\alpha_1\xi_i)}{\xi_{m+i}-\xi_i},$$

то мы получаем

$$a(i)-\alpha_1=\frac{(\varepsilon_{m+i}-\alpha_1\delta_{m+i})-(\varepsilon_i-\alpha_1\delta_i)}{\xi_{m+i}-\xi_i}.$$

Так как знаменатель $\xi_{m+i}-\xi_i$ положителен, то вероятность неравенства $a(i)-\alpha_1>0$ равна 1/2. Отсюда следует, что вероятность того, что точно j из $(a(i)-\alpha_1)$ больше нуля, т. е. $a(i)>\alpha_1$,

дается биномиальным распределением и равна $\binom{m}{j}\frac{1}{2^m}$, так что вероятность того, что r наибольших $a(i)$ больше α_1 и r наименьших $a(i)$ меньше α_1, равна *)

$$\mathsf{P}\{a(r) < \alpha_1 < a(m-r+1)\} = 1 - 2\sum_{j=0}^{r}\binom{m}{j}\frac{1}{2^m}. \qquad (29.116)$$

Последняя вероятность в соответствии с 5.7 может быть выражена в терминах неполной бета-функции. (29.116) дает доверительный интервал для α_1.

29.44 Делая дополнительное предположение, что δ и ε имеют нулевые медианы, мы получаем

$$\mathsf{P}\{\eta_i - \alpha_1\xi_i > \alpha_0\} = 1/2.$$

При любом заданном α_1 можно расположить значения $z_i = \eta_i - \alpha_1\xi_i$ в порядке возрастания и, как и в предыдущем случае, получить **)

$$\mathsf{P}\{z_r < \alpha_0 < z_{m-r+1}\} = 1 - 2\sum_{j=0}^{r}\binom{m}{j}\frac{1}{2^m}. \qquad (29.117)$$

С помощью этого метода не удается найти совместных доверительных интервалов для α_0 и α_1, за исключением таких, для которых известна лишь верхняя граница коэффициента доверия (см. упражнение 29.10). В упражнении 29.9 предложен критерий линейности.

Использование всех пар и применение (29.115) более сложно, поскольку распределения уже не являются биномиальными. В действительности они совпадают с распределениями, возникающими при получении распределения рангового коэффициента корреляции t, рассматриваемого в главе 31. Зная эти распределения, доверительные интервалы можно строить прежним способом.

29.45 Рассмотренные методы могут быть обобщены на случай линейной зависимости от k переменных. Если мы можем разбить n наблюдений на k групп, порядок которых будет одним и тем же как для наблюдаемых, так и для ненаблюдаемых значений одной из переменных, то, найдя центр тяжести каждой из групп, мы сможем построить гиперплоскость, проходящую через k точек. Если, кроме того, порядок наблюдаемых значений

*) В (29.116) $a(r)$ и $a(m-r+1)$ обозначают r-е и $(m-r+1)$-е по величине значения $a(i)$ в отличие от предыдущего текста, где номер i был связан с порядком значений ξ_i, а не $a(i)$. (*Прим. ред.*)

**) Аналогично предыдущему примечанию, индекс при z в (29.117) относится к упорядочению значений z по возрастанию. (*Прим. ред.*)

совпадает с порядком соответствующих ненаблюдаемых значений во всех точках, то можно получить $[n/k] = l$ соотношений для точек $(\xi_1, \xi_{l+1}, \xi_{2l+1}, \ldots, \xi_{kl+1})$, $(\xi_2, \xi_{l+2}, \xi_{2l+2}, \ldots, \xi_{kl+2})$ и т. д. и осреднить их. В принципе можно также обобщить и метод, использующий (29.115), однако с практической точки зрения вычисление всех возможных $\binom{n}{k}$ соотношений представляется слишком громоздким. См. упражнение 29.10.

29.46 Более радикальным методом является использование рангов значений ξ, т. е. натуральных чисел от 1 до n, в качестве инструментальной переменной. Этот метод должен быть эффективнее методов группировки, так как он использует больше информации. Дорфф и Гёрлэнд (1961b) показали, что он в общем случае более эффективен при малых объемах выборок, чем методы двух и трех групп, имея меньшее смещение и среднеквадратичную ошибку. Мы проиллюстрируем этот метод на примере.

Пример 29.7

Для данных примера 29.1 возьмем в качестве значений инструментальной переменной ζ ранги ξ от 1 до 9. Поскольку значения ξ уже упорядочены, мы просто пронумеруем их от 1 до 9. Получаем

$$\sum_i \zeta_i \eta_i = (1 \times 6{,}9) + (2 \times 12{,}5) + \ldots + (9 \times 39{,}1) = 1267{,}7,$$

$$\sum_i \zeta_i \xi_i = (1 \times 1{,}8) + (2 \times 4{,}1) + \ldots + (9 \times 18{,}9) = 549{,}0.$$

Согласно ранее сделанным вычислениям

$$\bar{\eta} = 23{,}14, \qquad \bar{\xi} = 9{,}57.$$

Кроме того, $\sum_i \zeta_i = \frac{1}{2} n(n+1) = 45$, так что ковариации *относительно выборочных средних* будут равны

$$\sum_i \zeta_i \eta_i - \bar{\eta} \sum_i \zeta_i = 1267{,}7 - 23{,}14 \times 45 = 226{,}40,$$

$$\sum_i \zeta_i \xi_i - \bar{\xi} \sum_i \zeta_i = 549{,}0 - 9{,}57 \times 45 = 118{,}35.$$

Отсюда, учитывая (29.91), находим

$$a_1 = \frac{226{,}40}{118{,}55} = 1{,}91.$$

Это же значение было получено методом двух групп в примере 29.4. Оно ближе к истинному значению 2, чем значение 1,86, полученное в примере 29.5 методом трех групп.

Контролируемые переменные

29.47 Берксон (1950) (см. также Линдли (1953b)) привел одно соображение, показывающее, что в некоторых типах экспериментов оценивание линейной связи между двумя переменными может быть сведено к регрессионной задаче. Напомним (см. **29.4**)), что соотношение

$$y = \alpha_0 + \alpha_1\xi + (\varepsilon - \alpha_1\delta)$$

не может рассматриваться как обычная регрессия, поскольку случайная величина ξ коррелирована с $(\varepsilon - \alpha_1\delta)$.

Пусть теперь проводится эксперимент, имеющий целью определение зависимости между y и x. Предположим, что в этом эксперименте мы можем заставить ξ принять любое заданное значение и затем измерить соответствующее значение η. Например, при определении зависимости между растяжением и натяжением пружины мы можем подвешивать грузики весом (ξ) 10 *г*, 20 *г*, 30 *г*, ..., 100 *г*, и измерять растяжения (η), которые можно рассматривать как результат влияния случайной ошибки ε на истинное значение y. Однако наши веса тоже могли быть измерены неточно. Так, считая вес равным 50 *г*, мы рискуем сделать это с ошибкой $\delta = 50 - x$. Повторяя эксперимент несколько раз с весом, принимаемым за 50 *г*, мы в действительности проводим эксперименты с истинными весами x_i, измеренными с ошибками $\delta_i = 50 - x_i$. Таким образом, используемые веса равны значениям случайной величины x. ξ называется контролируемой переменной, так как ее значение фиксировано заранее, тогда как неизвестное истинное значение случайно колеблется.

Пусть ошибки δ имеют нулевое среднее. Тогда среднее x будет равно $50 = \xi$. Таким образом, $\xi_i = x_i + \delta_i$, причем коэффициент корреляции между x_i и δ_i равен -1. Если предположить, как и прежде, что δ_i имеют одно и то же распределение для всех ξ_i, то можно написать $x = \xi - \delta$ и

$$\eta = (\alpha_0 + \alpha_1 x) + \varepsilon. \tag{29.118}$$

Подставляя в (29.118) предыдущее соотношение, находим

$$\eta = \alpha_0 + \alpha_1\xi + (\varepsilon - \alpha_1\delta). \tag{29.119}$$

Соотношение (29.119) имеет вид (29.8), однако существенно отличается от последнего. Поскольку ξ теперь уже не является случайной величиной, то ни ε, ни δ не коррелированы с ξ. Таким образом, (29.119) представляет собой обычное уравнение регрессии, к которому без всяких изменений можно применить обычные методы наименьших квадратов. Оценивание α_0 и α_1 и проверка гипотез, касающихся этих параметров, в данном случае не вызывают трудностей.

29.48 Даже если значения, в которых контролируются ξ, сами являются случайными величинами (т. е. определяемыми с помощью некоторого процесса случайного выбора), полученные выше выводы остаются в силе, если δ и ε некорреллированы с ξ. Предположение относительно ε обычно выполнено, но с δ дело может обстоять сложнее. Обращаясь к предыдущему примеру, предположим, что мы случайным образом выбираем значения весов для эксперимента. Требование некоррелированности δ и ξ в данном случае означает, например, что с бóльшими весами не должна быть связана тенденция к бóльшим или меньшим ошибкам в определении истинного веса по сравнению с малыми весами. Выполнение этого условия может быть проверено только эмпирическим путем.

Не вызывает сомнения, что во многих экспериментальных ситуациях рассмотренная модель применима. ξ часто представляет собой показание измерительного прибора, и экспериментатор обычно старается устанавливать ξ на некоторых заранее заданных значениях (выбранных не случайно, а с намерением по возможности полно отразить некоторый диапазон изменения). Поступая таким образом, он хорошо осознает, что прибор подвержен ошибкам и не показывает точно истинные значения x. Очень приятно обнаружить, после предшествующего обсуждения трудностей, возникающих в случаях, когда измерения производятся с ошибками, что в такой распространенной экспериментальной ситуации может быть использован стандартный метод НК. Этот пример иллюстрирует общее положение, важность которого трудно переоценить: тщательный анализ источников ошибок и структуры процесса наблюдений полезен для выбора наиболее подходящего статистического метода и может, как в данном случае, привести к простому решению на первый взгляд трудной задачи.

29.49 Анализ более сложных ситуаций, когда некоторые переменные контролируемы, а некоторые — нет или когда имеются повторные наблюдения при некоторых значениях контролируемых переменных, требует тщательного уточнения рассматриваемой модели. У нас нет возможности рассматривать возникающие здесь сложности. Интересные работы в этом отношении принадлежат Андерсону (1955) и Шеффе (1958).

Криволинейные зависимости

29.50 До сих пор рассматривались только линейные соотношения между переменными. Обобщение полученных методов на случай криволинейной зависимости не проходит так же непосредственно, как в теории регрессии, а отдельные задачи до сих пор не решены. Мы изложим здесь некоторые относящиеся

к этому вопросу результаты, в основном принадлежащие Гири (1942b, 1943, 1949, 1953).

Можно понять, какого рода трудности здесь возникают, если рассмотреть квадратичную функциональную зависимость

$$Y = \alpha_0 + \alpha_1 X + \alpha_2 X^2. \tag{29.120}$$

При тех же предположениях относительно ошибок ε и δ, что и в линейном случае, и при дополнительном упрощении, что отношение их дисперсий σ_δ^2 и σ_ε^2 известно (без ограничения общности мы будем считать его равным 1), получим в случае нормально распределенных ошибок следующее выражение для функции правдоподобия:

$$\log L = \text{const} - 2n \log \sigma_\varepsilon - \\ - \frac{1}{2\sigma_\varepsilon^2}\left\{\sum_i (\xi_i - X_i)^2 + \sum_i (\eta_i - \alpha_0 - \alpha_1 X_i - \alpha_2 X_i^2)^2\right\}. \tag{29.121}$$

Дифференцирование (29.121) дает

$$\xi_i - X_i + (\alpha_0 + \alpha_1 X_i + \alpha_2 X_i^2 - \eta_i)(\alpha_1 + 2\alpha_2 X_i) = 0, \tag{29.122}$$

$$i = 1, 2, \ldots, n,$$

$$\sum_i (\alpha_0 + \alpha_1 X_i + \alpha_2 X_i^2 - \eta_i) = 0, \tag{29.123}$$

$$\sum_i (\alpha_0 + \alpha_1 X_i + \alpha_2 X_i^2 - \eta_i) X_i = 0, \tag{29.124}$$

$$\sum_i (\alpha_0 + \alpha_1 X_i + \alpha_2 X_i^2 - \eta_i) X_i^2 = 0, \tag{29.125}$$

$$-\frac{2n}{\sigma_\varepsilon} + \frac{1}{\sigma_\varepsilon^3}\left[\sum_i (\xi_i - X_i)^2 + \sum_i (\eta_i - \alpha_0 - \alpha_1 X_i - \alpha_2 X_i^2)^2\right] = 0. \tag{29.126}$$

Суммируя (29.122) по i и используя (29.123) и (29.124), находим аналогично (29.42), что для значений ξ, измеренных относительно выборочного среднего, выполняется соотношение

$$\left(\sum_i \hat{X}_i\right) = \sum_i \xi_i = 0. \tag{29.127}$$

Если мы будем измерять значения η также относительно их среднего, то из (29.123)—(29.125) и (29.127) получим

$$\left.\begin{aligned} n\alpha_0 \qquad\qquad + \alpha_2 \sum X_i^2 &= 0, \\ \alpha_1 \sum X_i^2 + \alpha_2 \sum X_i^3 &= \sum \eta_i X_i, \\ \alpha_0 \sum X_i^2 + \alpha_1 \sum X_i^3 + \alpha_2 \sum X_i^4 &= \sum \eta_i X_i^2. \end{aligned}\right\} \tag{29.128}$$

Система (29.128) характерна для регрессионного анализа, однако здесь значения X ненаблюдаемы. Для получения МП-оце-

нок нам надо решить систему из $n+3$ уравнений (29.122) и (29.128) относительно $n+3$ неизвестных X_i, α_0, α_1, α_2. Оценка для σ_ε^2 получается затем из (29.126). На практике нам придется, вероятно, решать эту систему итерационными методами.

Усложнение задачи очевидно и с геометрической точки зрения. В соответствии с (29.121) требуется найти такую кривую второго порядка, чтобы сумма квадратов кратчайших расстояний от всех точек до этой кривой была минимальной. Отрезки, соединяющие точки с ближайшими точками кривой, не будут параллельными и даже могут определяться неоднозначно. Хотя численное решение может быть найдено, его трудно охарактеризовать наглядно.

29.51 Можно обобщить метод оценивания коэффициентов с помощью смешанных семиинвариантов. Рассмотрим кубическое структурное соотношение

$$y = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3. \tag{29.129}$$

От предположения нормальности ошибок мы отказываемся. Совместная х. ф. y и x имеет вид

$$\varphi(t_1, t_2) = \int \exp(\theta_1 y + \theta_2 x)\, dF(x, y),$$

где $\theta_j = it_j$. Учитывая (29.129), получаем

$$\left(\frac{\partial}{\partial\theta_1} - \alpha_0 - \alpha_1 \frac{\partial}{\partial\theta_2} - \alpha_2 \frac{\partial^2}{\partial\theta_2^2} - \alpha_3 \frac{\partial^3}{\partial\theta_2^3}\right)\varphi =$$

$$= \int (y - \alpha_0 - \alpha_1 x - \alpha_2 x^2 - \alpha_3 x^3) \exp(\theta_1 y + \theta_2 x)\, dF = 0. \tag{29.130}$$

Полагая $\psi = \log\varphi$ и используя соотношения

$$\left.\begin{aligned} \frac{\partial\varphi}{\partial\theta} &= \varphi \frac{\partial\psi}{\partial\theta}, \\ \frac{\partial^2\varphi}{\partial\theta^2} &= \varphi\left\{\frac{\partial^2\psi}{\partial\theta^2} + \left(\frac{\partial\psi^2}{\partial\theta}\right)^2\right\}, \\ \frac{\partial^3\varphi}{\partial\theta^3} &= \varphi\left\{\frac{\partial^3\psi}{\partial\theta^3} + 3\frac{\partial^2\psi}{\partial\theta^2}\cdot\frac{\partial\psi}{\partial\theta} + \left(\frac{\partial\psi}{\partial\theta}\right)^3\right\}, \end{aligned}\right\} \tag{29.131}$$

находим из (29.130)

$$\frac{\partial\psi}{\partial\theta_1} - \alpha_0 - \alpha_1\frac{\partial\psi}{\partial\theta_2} - \alpha_2\left\{\frac{\partial^2\psi}{\partial\theta^2} + \left(\frac{\partial\psi}{\partial\theta_2}\right)^2\right\} -$$

$$- \alpha_3\left\{\frac{\partial^3\psi}{\partial\theta_2^3} + 3\frac{\partial^2\varphi}{\partial\theta_2^2}\frac{\partial\varphi}{\partial\theta_2} + \left(\frac{\partial\psi}{\partial\theta_2}\right)^3\right\} = 0. \tag{29.132}$$

Разлагая ψ в степенной ряд и приравнивая нулю коэффициенты в (29.132), получаем систему уравнений для определения α_i,

$i = 0, 1, 2, 3$. Эти уравнения линейны относительно α_i, но в общем случае нелинейны относительно семиинвариантов. В силу (29.81) смешанные семиинварианты x и y совпадают с соответствующими семиинвариантами ξ и η. Если x и y распределены нормально, то метод становится неприменимым, как и в **29.30**.

29.52 При использовании изложенного метода возникают некоторые новые затруднения. Мы проиллюстрируем их на примере оценивания α_0, α_1 и α_2 в случае квадратичной зависимости ($\alpha_3 = 0$ в (29.129)).

Имеем

$$\psi = \sum \varkappa_{rs} \frac{\theta_1^r}{r!} \cdot \frac{\theta_2^s}{s!}.$$

Не теряя общности, можно положить $\varkappa_{10} = \varkappa_{01} = 0$, что эквивалентно выбору в качестве начала отсчета среднего значения x, за оценку которого берется среднее значение ξ. Из (29.132) получаем

$$\sum \varkappa_{rs} \frac{\theta_1^{r-1}\theta_2^s}{(r-1)!\,s!} - \alpha_0 - \alpha_1 \sum \varkappa_{rs} \frac{\theta_1^r\theta_2^{s-1}}{r!\,(s-1)!} - \\ - \alpha_2 \left\{ \sum \varkappa_{rs} \frac{\theta_1^r\theta_2^{s-2}}{r!\,(s-2)!} + \left(\sum \varkappa_{rs} \frac{\theta_1^r\theta_2^{s-1}}{r!\,(s-1)!} \right)^2 \right\} = 0. \qquad (29.133)$$

Приравнивая нулю постоянный коэффициент в (29.133), находим

$$-\alpha_0 - \alpha_2\varkappa_{02} = 0. \qquad (29.134)$$

Параметр α_0 входит только в это уравнение. Оно поэтому пригодно только для оценивания α_0, причем для этого требуется оценить не только α_2, но и семиинвариант $\varkappa_{02}$, зависящий от дисперсии ошибки. Далее получаем, приравнивая нулю следующие коэффициенты:

$$\left.\begin{aligned}
&\text{при } \theta_1 && \varkappa_{20} - \alpha_1\varkappa_{11} - \alpha_2\varkappa_{12} = 0,\\
&\text{при } \theta_2 && \varkappa_{11} - \alpha_1\varkappa_{02} - \alpha_2\varkappa_{03} = 0,\\
&\text{при } \theta_1^2 && \varkappa_{30} - \alpha_1\varkappa_{21} - \alpha_2(\varkappa_{22} + 2\varkappa_{11}^2) = 0,\\
&\text{при } \theta_1\theta_2 && \varkappa_{21} - \alpha_1\varkappa_{12} - \alpha_2(\varkappa_{13} + 2\varkappa_{11}\varkappa_{02}) = 0,\\
&\text{при } \theta_2^2 && \varkappa_{12} - \alpha_1\varkappa_{03} - \alpha_2(\varkappa_{04} + 2\varkappa_{02}^2) = 0.
\end{aligned}\right\} \qquad (29.135)$$

В первое из уравнений (29.135) входит семиинвариант $\varkappa_{20}$, который дальше не встречается. Поэтому это уравнение непригодно для оценки параметров α_i. Из (29.135) ясно, что коэффициент при любой степени θ_1, скажем при θ_1^r, содержит семиинвариант $\varkappa_{r+1,\,0}$, который не появляется ни в каком другом уравнении. Все эти уравнения поэтому бесполезны для оцени-

вания параметров α_i, если не предполагать, что ошибки ε распределены нормально (и, следовательно, имеют равные нулю семиинварианты порядка выше второго). В последнем случае семиинварианты $\varkappa_{r0}$ и смешанные семиинварианты могут быть оценены по выборке, что даст возможность использовать уравнения типа третьего в (29.135). В этом случае α_1 и α_2 находятся из третьего уравнения и уравнения, полученного в результате исключения $\varkappa_{02}$ из второго и четвертого уравнений в (29.135).

Если не предполагать нормальности ошибок, то нужны дальнейшие уравнения. Для последующих коэффициентов из (29.133) имеем:

$$\left.\begin{aligned} &\text{при } \theta_1^2\theta_2 \quad \varkappa_{31} - \alpha_1\varkappa_{22} - \alpha_2(\varkappa_{23} + 2\varkappa_{02}\varkappa_{21} + 4\varkappa_{11}\varkappa_{12}) = 0,\\ &\text{при } \theta_1\theta_2^2 \quad \varkappa_{22} - \alpha_1\varkappa_{13} - \alpha_2(\varkappa_{14} + 2\varkappa_{03}\varkappa_{11} + 4\varkappa_{02}\varkappa_{12}) = 0,\\ &\text{при } \theta_2^3 \qquad \varkappa_{13} - \alpha_1\varkappa_{04} - \alpha_2(\varkappa_{05} + 6\varkappa_{02}\varkappa_{03}) = 0. \end{aligned}\right\} \quad (29.136)$$

Два первых уравнения в (29.136) содержат, кроме смешанных семиинвариантов, также семиинварианты $\varkappa_{02}$ и $\varkappa_{03}$. Исключая последние с помощью второго и четвертого уравнений (29.135), получаем систему для нахождения α_1 и α_2. Оценивая затем $\varkappa_{02}$, мы сможем найти из (29.134) α_0. Некоторые из исключаемых членов могут быть нелинейными, что может привести к получению более чем одной системы оценок.

29.53 Метод двух групп, изложенный в **29.63**, очевидно, обобщается на случай полиномов k-го порядка, если удается разбить наблюдения на $k+1$ групп, порядок следования которых по ξ совпадает с порядком следования по x. В качестве решения берется полином, проходящий через центры тяжести этих групп. Метод Тейла (29.42) также легко обобщить. Разделив наблюдения на $k+1$ групп и беря по одной точке из каждой группы, мы сможем провести $n/(k+1)$ парабол, которые надо затем осреднить. Однако это не так просто, как кажется. Дело в том, что не обязательно лучшей оценкой истинной параболы будет парабола $\bar{a}_0 + \bar{a}_1 x + \bar{a}_2 x^2$, полученная путем усреднения коэффициентов. Можно еще указать некоторые эвристические способы усреднения, такие, как метод наименьших квадратов в направлении оси y или графическое изображение всех кривых и выбор из них той, которая, как кажется, лучше остальных представляет среднее положение на рассматриваемом участке.

29.54 При обобщении метода контролируемых переменных на нелинейный случай появляются некоторые особенности. Так, если линейный случай сводится к регрессионному анализу, то в нелинейном случае этого не удается сделать. Рассмотрим кубическую функциональную зависимость

$$Y = \alpha_0 + \alpha_1 X + \alpha_2 X^2 + \alpha_3 X^3,$$

Полагая $\xi = X + \delta$, $\eta = Y + \varepsilon$, получаем

$$\eta = \alpha_0 + \alpha_1(\xi - \delta) + \alpha_2(\xi - \delta)^2 + \alpha_3(\xi - \delta)^3, \qquad (29.137)$$

где контролируемые значения ξ фиксированы. Допустим, что для одного и того же набора значений ξ наблюдения повторяются. Обозначим $\mathbf{M}$ соответствующее математическое ожидание. Суммируя (29.137) по всем наблюдениям, получаем

$$\sum \eta = n\alpha_0 + \alpha_1 \sum \xi + \alpha_2 \sum \xi^2 + \alpha_2 \sum \delta^2 + \alpha_3 \sum \xi^3 + \\ + 3\alpha_3 \sum \xi\delta^2 + \text{члены нечетного порядка по } \delta \text{ и } \varepsilon.$$

Беря затем математические ожидания, находим

$$\mathbf{M}\left(\sum \eta\right) = \eta\left(\alpha_0 + \alpha_2\sigma_\delta^2\right) + \left(\alpha_1 + 3\alpha_3\sigma_\delta^2\right)\sum \xi + \alpha_2 \sum \xi^2 + \alpha_3 \sum \xi^3.$$

Аналогично, умножая (29.137) на ξ и суммируя, получаем систему уравнений

$$\mathbf{M}\left(\sum \eta\xi\right) = \left(\alpha_0 + \alpha_2\sigma_\delta^2\right)\sum \xi + \\ + \left(\alpha_1 + 3\alpha_3\sigma_\delta^2\right)\sum \xi^2 + \alpha_2 \sum \xi^3 + \alpha_3 \sum \xi^4 \qquad (29.138)$$

и т. д. Эти уравнения могут быть разрешены относительно величин

$$\left(\alpha_0 + \alpha_2\sigma_\delta^2\right), \quad \left(\alpha_1 + 3\alpha_3\sigma_\delta^2\right), \quad \alpha_2, \quad \alpha_3.$$

Интересно, что, хотя α_2 и α_3 идентифицируемы, α_0 и α_1 не являются таковыми, если не известна дисперсия σ_ε^2 или ее оценка. По-видимому, единственный способ обойти эту трудность состоит в повторении эксперимента при том же наборе значений ξ. Более полное изложение этого вопроса содержится в статьях Гири (1953) и Шеффе (1958).

29.55 Иногда нелинейное соотношение с помощью преобразования можно свести к линейному. Рассмотрим, например, функциональное соотношение

$$Y^\beta X^\gamma = \text{const}.$$

Здесь, очевидно, следует воспользоваться логарифмическим преобразованием. Однако эта процедура может иметь и нежелательные свойства. Так, если ошибки Y и X нормально распределены и гомоскедастичны, то ошибки преобразованных величин $\log Y$ и $\log X$ уже не будут обладать этими свойствами. Поэтому в данном случае, как, впрочем, и в любом другом, следует попытаться получить возможно больше априорной информации относительно ошибок наблюдений, а когда ошибки значительны и распределение их неизвестно, надо использовать методы оценивания, в которых число предположений о природе ошибок минимально.

Влияние ошибок наблюдения на регрессионный анализ

29.56 Нам кажется уместным завершить эту главу кратким рассмотрением вопроса о влиянии ошибок наблюдения на результаты регрессионного анализа. Предположим, что случайные величины x и y измеряются с ошибками, так что в действительности мы наблюдаем величины

$$\xi = x + \delta, \qquad \eta = y + \varepsilon.$$

Как и раньше, предположим, что ошибки δ и ε независимы между собой и независимы друг от друга. Возникает следующий вопрос. Пусть мы нашли регрессию η по ξ: как она связана с регрессией y по x, которая, собственно, нас и интересует?

Из **29.28** следует, что смешанные семиинварианты ξ и η совпадают с соответствующими семиинвариантами x и y. В частности, $\operatorname{cov}(\xi, \eta) = \operatorname{cov}(x, y)$. Однако коэффициенты регрессии зависят также от дисперсий, которые не остаются неизменными. Уравнение линейной регрессии

$$y = \beta_1 x, \quad \text{где} \quad \beta_1 = \operatorname{cov}(x, y)/\sigma_x^2,$$

заменяется уравнением

$$\eta = \beta_1' \xi, \quad \text{где} \quad \beta_1' = \operatorname{cov}(\xi, \eta)/\sigma_\xi^2 = \operatorname{cov}(x, y)/\{\sigma_x^2 + \sigma_\delta^2\}.$$

Поскольку $\beta_1' < \beta_1$, то влияние ошибок выражается в уменьшении наклона линии регрессии. Кроме того, уменьшается корреляция между ξ и η по сравнению с корреляцией между x и y.

29.57 Влияние ошибок, однако, не исчерпывается уменьшением указанных коэффициентов. Предположим, что истинная регрессия y по x в точности линейна и ошибки однородны (см. **28.7**). Следует ли отсюда, что регрессия η по ξ тоже в точности линейна? В общем случае ответ оказывается отрицательным, и только при некоторых довольно сильных условиях линейность сохраняется. Мы докажем одну изящно формулируемую теорему Линдли (1947): для сохранения линейности регрессии необходимо и достаточно, чтобы п. ф. с. случайной величины x была кратна п. ф. с. ошибки δ. Более точно, между п. ф. с. x и ξ должно выполняться соотношение

$$\beta_1 \psi_x = \beta_1' \psi_\xi. \tag{29.139}$$

В **28.7** было показано, что необходимым и достаточным условием точной линейности регрессии y по x вида $y = \beta_0 + \beta_1 x + \varepsilon$ с однородными ошибками служит возможность следующего представления совместной плотности x и y:

$$f(x, y) = g(x)\, h(y - \beta_0 - \beta_1 x), \tag{29.140}$$

где $g(x)$ — маргинальное распределение x, или эквивалентное условие в терминах п. ф. с.:

$$\psi(t_1, t_2) = \psi_g(t_1 + t_2\beta_1) + \psi_h(t_2) + it_2\beta_0. \qquad (29.141)$$

Мы знаем (28.5), что если $\zeta(t_1, t_2)$ — совместная п. ф. с. ξ и η и регрессия η по ξ — точная линейная регрессия вида $\eta = \beta_0' + \beta_1'\xi + \varepsilon$, то

$$\left[\frac{\partial}{\partial t_2}\zeta(t_1, t_2)\right]_{t_2=0} = i\beta_0' + \beta_1'\frac{\partial\zeta(t_1, 0)}{\partial t_1}. \qquad (29.142)$$

Однако если δ и ε независимы между собой и, кроме того, не зависят от x и y, то выполняется соотношение

$$\zeta(t_1, t_2) = \psi(t_1, t_2) + \psi_\delta(t_1) + \psi_\varepsilon(t_2). \qquad (29.143)$$

Подставляя (29.143) и (29.141) в (29.142), получаем

$$\left[\frac{\partial\psi_g(t_1 + t_2\beta_1)}{\partial t_2}\right]_{t_2=0} + \left[\frac{\partial\psi_h(t_2)}{\partial t_2}\right]_{t_2=0} + i\beta_0 + \left[\frac{\partial\psi_\varepsilon(t_2)}{\partial t_2}\right]_{t_2=0} =$$

$$= i\beta_0' + \beta_1'\frac{\partial\zeta(t_1, 0)}{\partial t_1}. \qquad (29.144)$$

Приравнивая коэффициенты при t_1 в (29.144), находим

$$\beta_1\frac{\partial\psi_g(t_1)}{\partial t_1} = \beta_1'\frac{\partial\zeta(t_1, 0)}{\partial t_1}. \qquad (29.145)$$

Интегрируя (29.145) и учитывая, что $\zeta(t_1, 0) = \psi_\xi(t_1)$, мы приходим к (29.139).

Из (29.144) также получаем

$$\mu_h + \beta_0 + \mu_\varepsilon = \beta_0', \qquad (29.146)$$

где μ_h и μ_ε обозначают соответствующие средние. В частности, если δ и ε имеют нулевые средние, то

$$\beta_0 = \beta_0'. \qquad (29.147)$$

Этим завершается доказательство необходимости (29.139). Доказательство достаточности не вызывает затруднений.

29.58 Если предположить, что и ошибки в регрессии η по ξ однородны, то можно получить гораздо более сильный результат (Кендалл, 1951, 1952). Действительно, в этом случае для п. ф. с. $\zeta(t_1, t_2)$ выполняется как (29.143), так и (29.141). Ставя штрихи в (29.141), получим

$$\zeta(t_1, t_2) = \psi(t_1, t_2) + \psi_\delta(t_1) + \psi_\varepsilon(t_2) =$$

$$= \psi_{g'}(t_1 + t_2\beta_1') + \psi_{h'}(t_2) + it_2\beta_0'. \qquad (29.148)$$

Подставляя $\psi(t_1, t_2)$ из (29.141) в (29.148), получаем

$$\psi_g(t_1 + t_2\beta_1) + \psi_h(t_2) + it_2\beta_0 + \psi_\delta(t_1) + \psi_\varepsilon(t_2) = \\ = \psi_{g'}(t_1 + t_2\beta_1') + \psi_{h'}(t_2) + it_2\beta_0'. \quad (29.149)$$

Полагая в (29.149) $t_1 = 0$ и вычитая полученное уравнение из (29.149), находим

$$\psi_g(t_1 + t_2\beta_1) - \psi_g(t_2\beta_1) + \psi_\delta(t_1) = \\ = \psi_{g'}(t_1 + t_2\beta_1') - \psi_{g'}(t_2\beta_1'). \quad (29.150)$$

Придавая семиинвариантам x и ξ верхние индексы g и g', из (29.150) получаем

$$\sum \varkappa_r^g \frac{(t_1 + t_2\beta_1)^r}{r!} - \sum \varkappa_r \frac{(t_2\beta_1)^r}{r!} + \sum \varkappa_r^\delta \frac{t_2^r}{r!} = \\ = \sum \varkappa_r^{g'} \frac{(t_1 + t_2\beta_1')^r}{r!} - \sum \varkappa_r^{g'} \frac{(t_2\beta_1')^r}{s!}. \quad (29.151)$$

Рассмотрим теперь член порядка $r > 2$, скажем порядка $r = 3$. Приравнивая коэффициенты, имеем

$$\left.\begin{aligned} \varkappa_3^g + \varkappa_3^\delta &= \varkappa_3^{g'}, \\ \beta_1\varkappa_3^g &= \beta_1'\varkappa_3^{g'}, \\ \beta_1^2\varkappa_3^g &= (\beta_1')^2\varkappa_3^{g'}. \end{aligned}\right\} \quad (29.152)$$

Второе и третье соотношения могут выполняться, только если $\beta_1 = \beta_1'$ или $\varkappa_3^g$ и $\varkappa_3^{g'}$ обращаются в нуль. Отсюда следует, что обращаются в нуль все семиинварианты третьего порядка и, аналогично, семиинварианты более высоких порядков. Обратное также верно. Таким образом, точная линейная регрессия с однородными ошибками $y = \beta_0 + \beta_1 x$ превращается в точную линейную регрессию с однородными ошибками $\eta = \beta_0' + \beta_1'\xi$ тогда и только тогда, когда x, δ и ξ распределены нормально.

29.59 Имеются также другие теоремы, относящиеся к этому вопросу. По-видимому, первый результат был получен Аллен (1938), которая при довольно ограничительных предположениях доказала, что если

$$\left.\begin{aligned} \xi &= lx + \delta, \\ \eta &= mx + \varepsilon, \end{aligned}\right\} \quad l,\ m \neq 0,$$

то нормальность x и δ является необходимым и достаточным условием для точной линейности регрессии η по ξ для всех l из замкнутого интервала. Некоторые из ее условий были ослаблены Фикс (1949а), которая требовала только конечности средних x, δ и ε и конечности дисперсии x или δ. Результаты Фикс

были затем обобщены Лахой (1956) на случай зависимых ошибок δ и ε.

Линдли (1947) доказал результат более общий, чем (29.139) (см. упражнения 29.12, 29.14). Результат **29.58** также может быть обобщен на случай нескольких переменных (см. упражнение 29.13).

УПРАЖНЕНИЯ

29.1 Показать, что МП-оценка $\hat{\alpha}_1$, определенная в (29.29), является монотонной функцией от λ.

29.2 Используя метод, изложенный в **29.28**, показать, что если (а) ошибки δ не зависят от x и (б) можно найти $k-1$ обращающихся в нуль смешанных семиинвариантов δ, то уравнения (29.85) можно использовать для оценки параметров α. В частности, эти условия выполнены, если δ имеют многомерное нормальное распределение.

(Гири, 1942b.)

29.3 Показать, что уравнения (29.85) остаются в силе, если семиинварианты $\varkappa$ заменить соответствующими моментами x, однако это свойство не выполняется для наблюдений ξ.

29.4 Прямыми вычислениями показать, что оценка (29.98) является состоятельной оценкой для α_1 и что, следовательно, также состоятельны оценки (29.102) для дисперсий ошибок.

(Вальд, 1940.)

29.5 Используя равенство (29.111), показать, что доверительная область для α_0 и α_1 представляет собой внутренность эллипса.

(Вальд, 1940.)

29.6 Пусть имеется n наблюдений $x_1, x_2, \ldots, x_n$ векторной случайной величины, компоненты которой независимы и имеют нормальные распределения с единичными дисперсиями. Требуется проверить допустимость соотношения

$$\alpha_0 + \boldsymbol{\alpha}'\boldsymbol{x} = 0, \tag{A}$$

где α_0 — скаляр, а $\boldsymbol{\alpha}'$ — транспонированный вектор-столбец $\boldsymbol{\alpha}$. Показать, что

$$n\varphi = \sum_{j=1}^{n} \frac{(\alpha_0 + \boldsymbol{\alpha}'\boldsymbol{x}_j)'(\alpha_0 + \boldsymbol{\alpha}'\boldsymbol{x}_j)}{\boldsymbol{\alpha}'\boldsymbol{\alpha}} \tag{B}$$

имеет распределение χ^2 с n степенями свободы. Обозначая $\boldsymbol{V}$ матрицу рассеяния наблюдений, показать, что огибающая (A) при условии, что $n\varphi$ в (B) равно константе χ_0^2, описывается уравнением

$$|\boldsymbol{V} - \varphi\boldsymbol{I}| + V_{ij}^* x_i x_j = 0,$$

где $\boldsymbol{I}$ — единичная матрица, а V_{ij}^* — алгебраическое дополнение элемента (i, j) матрицы $|\boldsymbol{V} - \varphi\boldsymbol{I}|$. Показать, что последнее соотношение может быть также записано в виде

$$1 + \boldsymbol{x}'(\boldsymbol{V} - \varphi\boldsymbol{I})^{-1}\boldsymbol{x} = 0.$$

(Р. Браун и Фередей, 1958.)

29.7 Пусть в предыдущем упражнении корни уравнения $|\boldsymbol{V} - \varphi \boldsymbol{I}| = 0$ равны $\varphi_1, \ldots, \varphi_k$, и пусть $\boldsymbol{V}$ — диагональная матрица с элементами φ_j, а $\boldsymbol{L}$ — ортогональная матрица порядка $(k \times k)$, определенная соотношением $\boldsymbol{VL} = \boldsymbol{L\Lambda}$. Показать, что после преобразования к новым переменным $\boldsymbol{y} = \boldsymbol{L'x}$ огибающая имеет уравнение $1 + \boldsymbol{y}'(\boldsymbol{\Lambda} - \varphi \boldsymbol{I})^{-1}\boldsymbol{y} = 0$, т. е.

$$\sum_{j=1}^{k} \frac{y_j^2}{\varphi_j - \varphi} + 1 = 0.$$

(Р. Браун и Фередей, 1958.)

29.8 В условиях предыдущих двух упражнений показать, что если вместо (A) в упражнении 29.7 выполняется соотношение $\beta_0 + \boldsymbol{\beta}'\boldsymbol{y} = 0$, то совместная доверительная область для коэффициентов β имеет вид

$$\beta_0^2 + \sum_{j=1}^{k} (\varphi_j - \varphi_0)\beta_j^2 = 0,$$

где φ_0 — критическое значение статистики φ, определенной в (B), которое находится из таблиц χ^2-распределения.

(Р. Браун и Фередей, 1958.)

29.9 Показать, что статистики $a(i)$, определяемые равенством (29.114), можно использовать для построения критерия линейности зависимости (против выпуклой или вогнутой зависимости), рассматривая корреляцию между $a(i)$ и i.

(Тейл, 1950.)

29.10 Переменные X_j связаны с Y соотношением

$$Y = \alpha_0 + \sum_{j=1}^{k} \alpha_j X_j.$$

Ошибки наблюдения иксов таковы, что наблюденные $\xi_j (= X_j + \delta_j)$ расположены в том же порядке, что и X_j для всех $j = 1, 2, \ldots, k$. Показать, что если δ_j таковы, что

$$\mathrm{P}\{z_i < z_j\} = \frac{1}{2},$$

где $z_i = \varepsilon_i - \sum_{j=1}^{k} \alpha_j \delta_j$, то для каждого α_j при заданных остальных α можно построить доверительный интервал тем же способом, как в **29.43**. Исходя из этого, показать, как можно строить консервативную доверительную область для всех α, беря объединение отдельных интервалов.

(Тейл, 1950.)

29.11 Имеется $n (= 2l + 1)$ наблюдений величины $\eta (= Y + \varepsilon)$, сделанных в равноотстоящих с шагом 1 точках X. Величины X измерены без ошибки, т. е. $\xi = X$. Дисперсия ошибок ε равна σ^2. Показать, что если параметры в соотношении $Y = \alpha_0 + \alpha_1 X$ оценены по методу наименьших квадратов, дающему несмещенные оценки с минимальной дисперсией, то дисперсия оценки для α_1 равна $3\sigma^2/\{l(l+1)(2l+1)\}$.

Показать далее, что если наблюдения разделены на три группы, состоящие из nk, $n - 2nk$, nk наблюдений, то оценка (29.112) имеет максимальную эффективность, превышающую 8/9, когда $k = 1/3$, тогда как для $k = 1/2$ эффективность составляет лишь 27/32 от максимальной.

(Бартлетт, 1949.)

29.12 Пусть регрессия y по $x_1, x_2, \ldots, x_k$ дается соотношением $y = \sum_j \beta_j x_j$ и ошибки не зависят от x. Показать, что если x_j наблюдаются с ошибками δ_j ($\xi_j = x_j + \delta_j$), а y — с ошибкой ε ($\eta = y + \varepsilon$), причем ошибки δ не зависят друг от друга и от ε, то регрессия η по ξ в точности линейна и имеет вид

$$\eta = \sum_j \beta'_j \xi_j$$

тогда и только тогда, когда

$$\sum_j (\beta_j - \beta'_j) \frac{\partial \psi_{x_j}}{\partial t_j} = \sum_j \beta'_j \frac{\partial \psi_{\delta_j}}{\partial t_j},$$

где ψ — п. ф. с. случайной величины, указываемой индексом. Этот результат является обобщением **29.57.**

(Линдли, 1947.)

29.13 Продолжая упражнение 29.12, показать, что ошибки во второй регрессии не зависят от ξ тогда и только тогда, когда x, ξ и δ имеют нормальное распределение. Этот результат является обобщением **29.58.**

(Кендалл, 1951, 1952.)

29.14 Показать, что в упражнении 29.12

$$\sum_j \beta_j \frac{\partial^2 \psi_x}{\partial t_i \, \partial t_j} = \sum \beta'_j \frac{\partial^2 \psi_\xi}{\partial t_i \, \partial t_j},$$

откуда следует, что если $\boldsymbol{S}$ — матрица рассеяния ξ, а диагональная матрица $\boldsymbol{\Lambda}$ — матрица рассеяния δ, то

$$\boldsymbol{\beta} - (\boldsymbol{\beta}') = \boldsymbol{\beta} \boldsymbol{\Lambda} \boldsymbol{S}^{-1},$$

где $\boldsymbol{\beta}$, $(\boldsymbol{\beta}')$ — векторы-строки с компонентами β_j и β'_j соответственно.

(Линдли, 1947.)

29.15 Показать, что если ненаблюдаемые переменные x, y и инструментальная переменная ζ имеют нормальные распределения с нулевыми средними и a_1 определено в (29.91), то статистика

$$t^2 = (n-1) \Big/ \left\{ \frac{\sigma_\zeta^2 \left(a_1^2 \sigma_x^2 - 2a_1 \sigma_{xy} + \sigma_y^2\right)}{(a_1 \sigma_{x\zeta} - \sigma_{y\zeta})^2} - 1 \right\}$$

имеет распределение Стьюдента t^2 с $n-1$ степенями свободы.

(Гири, 1949.)

29.16 Пользуясь результатом последнего упражнения, показать, что приближенно

$$2n^2 \mathrm{D} a_1 = \left[\frac{\sigma_{x\zeta} (a_1 \sigma_{x\zeta} - \sigma_{y\zeta})}{\sigma_\zeta^2 \left(a_1^2 \sigma_x^2 - 2a_1 \sigma_{xy} + \sigma_y^2\right) - (a_1 \sigma_{x\zeta} - \sigma_{y\zeta})^2} - \right.$$
$$\left. - \frac{(a_1 \sigma_{x\zeta} - \sigma_{y\zeta})^2 \left\{ a_1 \left(\sigma_x^2 \sigma_\zeta^2 - \sigma_{x\zeta}^2\right) + \left(\sigma_{x\zeta} \sigma_{y\zeta} - \sigma_{xy} \sigma_\zeta^2\right) \right\}}{\left\{ \sigma_\zeta^2 \left(a_1^2 \sigma_x^2 - 2\sigma_1 \sigma_{xy} + \sigma_y^2\right) - (a_1 \sigma_{x\zeta} - \sigma_{y\zeta})^2 \right\}^2} \right]^{-2}.$$

(Мадански, 1959.)

29.17 Предполагая, что x и y наблюдаются без ошибок (т. е. $\xi = x$, $\eta = y$), показать, что для фиксированных x и ζ эффективность оценки (29.91) по сравнению с НК-оценкой равна квадрату коэффициента корреляции между x и ζ.

(Дёрбин, 1954.)

29.18 Пусть условия $s_\xi^2 > \sigma_\delta^2$, $s_\eta^2 \geqslant s_{\xi\eta}^2/(s_\xi^2 - \sigma_\delta^2)$ в случае 1 в **29.12** не выполнены. Показать, что ФП имеет максимумы при $\sigma_\varepsilon^2 = 0$ и при $\sigma_x^2 = 0$, и, сравнивая их, проверить, что глобальный максимум всегда достигается в первом случае. Вывести отсюда, что

$$\hat{\alpha}_1 = s_\eta^2/s_{\xi\eta}, \qquad \hat{\sigma}_\varepsilon^2 = 0, \qquad \hat{\sigma}_x^2 = s_{\xi\eta}^2/s_\eta^2.$$

(Бёрч, 1964а.)

ГЛАВА 30

КРИТЕРИИ СОГЛАСИЯ

30.1 При рассмотрении процедур оценки и проверки гипотез, начиная с главы 17, мы занимались исключительно задачами о параметрах распределений известной функциональной формы. В нашей классификации задач проверки гипотез в **22.3** мы дали определение непараметрической гипотезы, но до сих пор еще не рассматривали непараметрических задач проверки гипотез или оценивания. Группа из четырех глав, начиная с этой, будет посвящена систематическому изучению этого предмета.

Нам будет удобно отложить общее обсуждение непараметрических задач и их особенностей до главы 31. В настоящей главе мы ограничимся одним специальным классом процедур, который стоит несколько в стороне от других и имеет достаточно важное практическое значение, чтобы такое выделение было оправдано.

Критерии согласия

30.2 Пусть $x_1, x_2, \ldots, x_n$ — независимые наблюдения случайной величины с функцией распределения $F(x)$, которая нам неизвестна. Предположим, что нам нужно проверить гипотезу

$$H_0\colon\ F(x) = F_0(x), \tag{30.1}$$

где $F_0(x)$ — некоторая заданная ф. р., которая может быть непрерывной или дискретной. Задача проверки гипотезы (30.1) называется *задачей проверки согласия.* Любой критерий для (30.1) называется *критерием согласия.*

Гипотезы согласия, подобно параметрическим гипотезам, естественным образом разделяются на простые и сложные. Гипотеза (30.1) проста, если $F_0(x)$ полностью определена; например, гипотеза (а), что n наблюдений получены из нормального распределения с заданными средним и дисперсией, есть простая гипотеза. С другой стороны, если мы желаем проверить (б), что наблюдения получены из нормального распределения, параметры которого могут иметь любые значения, то эта гипотеза будет сложной (в этом случае часто говорят о «проверке нор-

мальности»). Аналогично, если (в) нормальное распределение имеет фиксированное среднее, но не дисперсию, гипотеза остается сложной. Это различие в точности того же характера, что и обсуждавшееся нами в параметрическом случае в **22.4**.

30.3 Ясно, что (30.1) — не более чем повторение общей постановки задачи проверки гипотез; мы всего лишь выразили гипотезу через ф. р. вместо плотности. Какой в этом смысл? Не собираемся ли мы повторить уже пройденный путь?

Имеются несколько причин для новой формулировки. Методы параметрических гипотез, развитые ранее, были по необходимости связаны с гипотезами, состоящими в наложении одного или нескольких ограничений (см. **22.4**) в пространстве параметров; они не дают никаких средств для проверки такой гипотезы, как (б) в **30.2**, где не налагается никаких ограничений на параметры и мы проверяем непараметрическую гипотезу, что функция распределения генеральной совокупности принадлежит к заданному (бесконечному) семейству распределений. В таких случаях и даже в тех случаях, когда гипотеза налагает одно или больше ограничений, как в (а) или (в) в **30.2**, новая формулировка гипотезы в форме (30.1) приводит нас к новым методам. Дело в том, что интуиция заставляет нас ожидать, что распределение совокупности выборочных значений в целом будет близко следовать истинной ф. р. $F(x)$. Естественно поэтому попытаться использовать выборочное распределение как средство для проверки (30.1), и мы увидим, что в наиболее важных критериях согласия именно это и делается. Кроме того, «оптимальные» критерии, которые мы строили для параметрических гипотез H_0, определялись свойствами их функций мощности против альтернативных гипотез, отличавшихся от H_0 только значениями параметров, фиксированных при H_0. Кажется по меньшей мере правдоподобным, что критерий, основанный на выборочном распределении, будет иметь удовлетворительные свойства мощности против более широкого (бесконечного) класса альтернатив, хотя он может и не быть оптимальным против какой бы то ни было из них.

Критерии отношения правдоподобия и Пирсона для простой гипотезы H_0

30.4 Эти два известных метода проверки согласия используют очень простой прием. Мы рассмотрим его сначала в том случае, когда $F_0(x)$ полностью определена, так что гипотеза (30.1) проста.

Предположим, что область значений величины x произвольным образом разбита на k взаимно непересекающихся классов. (На практике они обычно выбираются как последовательные

интервалы в области значений x, хотя это и не обязательно*)). Тогда, поскольку $F_0(x)$ известна, мы можем вычислить вероятности попадания наблюдения в каждый из классов. Если обозначить их p_{0i}, $i = 1, 2, \ldots, k$, а наблюденные частоты в k классах через n_i $\left(\sum_{i=1}^{k} n_i = n\right)$, то n_i имеют мультиномиальное распределение (см. **5.30**), и из (5.78) мы видим, что функция правдоподобия равна

$$L(n_1, n_2, \ldots, n_k \mid p_{01}, p_{02}, \ldots, p_{0k}) \propto \prod_{i=1}^{k} p_{0i}^{n_i}. \qquad (30.2)$$

С другой стороны, если истинная функция распределения есть $F_1(x)$, где F_1 может быть любой ф. р., мы можем обозначить вероятности в k классах через p_{1i}, $i = 1, 2, \ldots, k$, и ФП тогда будет равна

$$L(n_1, n_2, \ldots, n_k \mid p_{11}, p_{12}, \ldots, p_{1k}) \propto \prod_{i=1}^{k} p_{1i}^{n_i}. \qquad (30.3)$$

Теперь мы легко можем найти критерий отношения правдоподобия для гипотезы (30.1) при сложной альтернативной гипотезе

$$H_1\colon F(x) = F_1(x).$$

Функция правдоподобия (30.3) достигает максимума, когда мы подставляем вместо p_{1i} МП-оценки $\hat{p}_{1i} = n_i/n$. Статистикой ОП для проверки H_0 против H_1 будет, следовательно,

$$l = \frac{L(n_1, n_2, \ldots, n_k \mid p_{01}, p_{02}, \ldots, p_{0k})}{L(n_1, n_2, \ldots, n_k \mid \hat{p}_{11}, \hat{p}_{12}, \ldots, \hat{p}_{1k})} = n^n \prod_{i=1}^{k} (p_{0i}/n_i)^{n_i}. \qquad (30.4)$$

Гипотеза H_0 отвергается, если значение l достаточно мало. Точное распределение статистики (30.4) неизвестно. Однако мы знаем из **24.7**, что если верна H_0, то при $n \to \infty$ величина $-2 \log l$ распределена асимптотически как χ^2 с $k - 1$ степенями свободы (так как имеется $r = k - 1$ независимых ограничений p_{1i}, поскольку $\sum_{i=1}^{k} p_{1i} = 1$).

30.5 Однако (30.4) — это не та классическая статистика критерия, которая была предложена Карлом Пирсоном (1900) для этой ситуации. В этой процедуре, которая была уже получена в примере 15.3, используется асимптотическая k-мерная нормальность мультиномиального распределения величин n_i и тот факт, что при гипотезе H_0 квадратичная форма в показателе

*) Выбор k и классов будет обсуждаться ниже, в **30.20—23, 30.28—30.** Сейчас мы считаем их произвольными.

этого нормального распределения имеет распределение χ^2 с числом степеней свободы, равным ее рангу $k-1$. В примере 15.3 мы нашли, что эта квадратичная форма в используемых теперь обозначениях равна *)

$$X^2=\sum_{i=1}^{k}\frac{(n_i-np_{0i})^2}{np_{0i}}. \tag{30.5}$$

Уайз (1963, 1964) исследовал ошибку приближения, возникающего, когда (30.5) считают распределенной как χ^2_{k-1}, и показал, что ошибка особенно мала, когда np_{0i} равны или близки между собой, — тогда можно не требовать, чтобы они были большими (см. **30.22**, **30.30** для сложной гипотезы H_0).

Из (30.4) мы имеем

$$-2\log l=2\sum_{i=1}^{k}n_i\log(n_i/np_{0i}). \tag{30.6}$$

Таким образом, эти две различные статистики (30.5) и (30.6) имеют асимптотически одинаковое распределение при H_0. Более того, когда верна H_0, эти статистики асимптотически эквивалентны, так как если мы обозначим $\Delta_i=\frac{n_i-np_{0i}}{np_{0i}}$, то

$$-2\log l=2\sum_i n_i\log(1+\Delta_i)=$$

$$=2\sum_i\{(n_i-np_{0i})+np_{0i}\}\left\{\Delta_i-\tfrac{1}{2}\Delta_i^2+O(n^{-3/2})\right\}=$$

$$=2\sum_i\left\{(n_i-np_{0i})\Delta_i+np_{0i}\Delta_i-\frac{np_{0i}}{2}\Delta_i^2+O(n^{-1/2})\right\},$$

и поскольку $\sum_i p_{0i}\Delta_i=0$, то

$$-2\log l=\sum_i\{np_{0i}\Delta_i^2+O(n^{-1/2})\}=X^2\{1+O(n^{-1/2})\}. \tag{30.7}$$

При малых n эти статистики различаются. Статистика Пирсона (30.5) может быть преобразована к виду

$$X^2=\frac{1}{n}\sum_i\frac{n_i^2}{p_{0i}}-n, \tag{30.8}$$

*) Следуя практике последнего времени, мы обозначаем через X^2 статистику критерия и оставляем символ χ^2 для обозначения распределения, о котором мы так часто упоминали. Прежние авторы пользовались одним и тем же обозначением χ^2 как для статистики, так и для распределения.

более простому для вычислений; но запись (30.5) имеет перед (30.8) то преимущество, что в ней явно фигурируют разности между наблюденными частотами n_i и их гипотетическими ожиданиями np_{0i}, и сами эти разности, очевидно, представляют интерес. Соответствующее упрощение (30.6) неудобно для вычислений.

Выбор критической области

30.6 Поскольку H_0 отвергается при малых значениях l, из (30.7) следует, что при использовании (30.5) в качестве статистики критерия гипотеза H_0 должна отвергаться, когда X^2 велико. По этому вопросу в литературе не было полной уверенности, и более ранняя практика состояла в том, чтобы отвергать H_0 как при больших, так и при малых значениях X^2, т. е. применять не односторонний, а двусторонний критерий. Например, Кокрэн (1952) поддерживает эту практику на том основании, что причиной чрезвычайно малых значений X^2 являются скорее всего численные ошибки в вычислениях; или в других случаях такие значения возникают из-за того, что частоты n_i смещены, возможно не намеренно, с тем, чтобы сблизить их с гипотетическими ожиданиями np_{0i}.

Нет сомнения, что точность вычислений следует проверять, но это можно, по-видимому, делать более прямыми и эффективными методами, чем исследование полученного значения X^2. В конце концов, нет никакой уверенности в том, что умеренное и приемлемое значение X^2 было вычислено с большей точностью, чем очень малое. Второе соображение Кокрэна более убедительно, но ясно, что в этом случае мы рассматриваем другую и более редкую гипотезу (что сбор наблюдений подвержен умышленной или неумышленной нерегулярности), которую нужно точно сформулировать, прежде чем мы можем определить наилучшую критическую область для ее проверки (см. Стьюарт (1954а)). Оставляя в стороне такие нерегулярности, мы используем в качестве критической область больших значений X^2. Это будет оправдано с точки зрения асимптотической мощности в **30.27**.

30.7 Существенным моментом в критериях максимального правдоподия и Пирсона является переход к задаче о мультиномиальном распределении. Необходимость группировки данных по классам приводит к потере некоторого количества информации, особенно если наблюдаемая величина непрерывна. Однако с этим недостатком связано и соответствующее достоинство: нам нет необходимости знать значения индивидуальных наблюдений, как только мы имеем k классов, для которых могут быть вычислены гипотетические p_{0i}. На самом деле вовсе не

обязательно должна наблюдаться какая-то случайная величина — мы можем применять любой из двух критериев согласия, даже если исходные данные относятся к не числовой классификации. Это иллюстрируется примером 30.1.

Пример 30.1

В некоторых классических экспериментах с селекцией гороха Мендель наблюдал частоты различных видов семян, получаемых при скрещивании растений с круглыми желтыми семенами и растений с морщинистыми зелеными семенами. Они приводятся ниже вместе с теоретическими вероятностями по теории наследственности Менделя.

Семена	Наблюденная частота n_i	Теоретическая вероятность p_{0i}
Круглые и желтые	315	9/16
Морщинистые и желтые	101	3/16
Круглые и зеленые	108	3/16
Морщинистые и зеленые	32	1/16
	$n = 556$	1

Формула (30.8) дает

$$X^2 = \frac{1}{556} \cdot 16 \left\{ \frac{315^2}{9} + \frac{101^2}{3} + \frac{108^2}{3} + \frac{32^2}{1} \right\} - 556 =$$

$$= \frac{16}{556} \cdot 19\,337{,}3 - 556 = 0{,}47.$$

При $(k-1) = 3$ степенях свободы таблицы распределения χ^2 показывают, что вероятность значения, превышающего 0,47, лежит между 0,90 и 0,95, так что действительно между наблюдениями и теорией имеется очень хорошее согласие: критерий любого размера $\alpha \leqslant 0{,}90$ не отвергал бы эту гипотезу.

Для статистики максимального правдоподобия формула (30.6) дает после заметно бо́льших вычислений значение $-2 \log l = 0{,}48$, очень близкое к значению X^2.

Сложная гипотеза H_0

30.8 Ограничиваясь теперь статистикой критерия Пирсона (30.5), мы рассмотрим ситуацию, возникающую, когда проверяемая гипотеза сложна — критерий ОП остается асимптотически эквивалентным, когда верна H_0 (см. упражнение 30.11).

Предположим, что известна функциональная форма $F_0(x)$, но что некоторые (или, возможно, все) параметры остаются неизвестными, как в (б) или (в) пункта **30.2**. Новым в мультиномиальной формулировке пункта **30.4** является теперь то, что теоретические вероятности p_{0i} не поддаются непосредственному вычислению, так как они зависят от s ($< k-1$ по предположению) неизвестных параметров $\theta_1, \theta_2, \ldots, \theta_s$, совокупность которых мы обозначим $\boldsymbol{\theta}$. Таким образом, мы должны записывать их как $p_{0i}(\boldsymbol{\theta})$. Чтобы двигаться дальше, мы должны оценить $\boldsymbol{\theta}$ некоторым вектором оценок $\boldsymbol{t}$ и использовать (30.5) в форме

$$X^2 = \sum_{i=1}^{k} \frac{\{n_i - np_{0i}(t)\}^2}{np_{0i}(t)}.$$

Ясно, что это меняет задачу о распределении статистики, так как теперь $p_{0i}(t)$ сами являются случайными величинами, и совсем не очевидно, что асимптотическое распределение X^2 будет иметь ту же форму, что и в случае простой гипотезы H_0. Действительно, член $n_i - np_{0i}(t)$ не обязательно имеет нулевое математическое ожидание. Мы можем тождественно переписать X^2 как

$$X^2 = \sum_{i=1}^{k} \frac{1}{np_{0i}(t)} [\{n_i - np_{0i}(\theta)\}^2 + n^2\{p_{0i}(t) - p_{0i}(\theta)\}^2 - \\ - 2n\{n_i - np_{0i}(\theta)\}\{p_{0i}(t) - p_{0i}(\theta)\}]. \quad (30.9)$$

Из теории мультиномиального распределения мы знаем, что асимптотически

$$n_i - np_{0i}(\theta) \sim cn^{1/2},$$

так что первый член в квадратных скобках в (30.9) имеет порядок n. Если мы имеем также

$$p_{0i}(t) - p_{0i}(\theta) = o(n^{1/2}), \quad (30.10)$$

то второй и третий члены будут меньшего порядка, чем n, и будут асимптотически пренебрежимы, так что асимптотическое поведение выражения (30.9) определяется тогда его первым членом. Однако и это выражение содержит еще случайную величину $np_{0i}(t)$ в знаменателе, но с той же степенью приближения мы можем заменить ее на $np_{0i}(\theta)$. Мы видим, таким образом, что если выполняется (30.10), то статистика (30.9) ведет себя асимптотически точно так же, как (30.5),— она имеет распределение χ^2 с $k-1$ степенями свободы. Однако, если $p_{0i}(t)$ «хорошо ведут себя» как функции от t, то их отличие от $p_{0i}(\theta)$

будет того же порядка, что и отличие t от θ. Тогда во всех практических ситуациях (30.10) требует, чтобы

$$t - \theta = o(n^{-1/2}). \tag{30.11}$$

Соотношение (30.11), как правило, не выполняется, так как мы обычно имеем оценки с дисперсиями и ковариациями порядка n^{-1}, и тогда

$$t - \theta = O(n^{-1/2}). \tag{30.12}$$

В этом «регулярном» случае, следовательно, приведенное выше рассуждение теряет силу. Но оно остается в силе в тех случаях, когда оценки имеют дисперсии порядка n^{-2}, что, как мы видели, характерно для оценок параметров, определяющих граничные точки области значений случайной величины (см. примеры 14.8, 14.13 и 32.11). В таких случаях, следовательно, для применения (30.8) не требуется никакой новой теории. Более распространенный случай, когда выполнено (30.12), требует дальнейшего исследования.

30.9 Чтобы упростить наше обсуждение, мы приведем сначала иной вывод асимптотического распределения (30.5) в случае простой гипотезы, принадлежащий Фишеру (1922a).

Предположим, что имеются k независимых пуассоновских случайных величин, из которых i-я имеет параметр np_{0i}. Вероятность того, что первая величина примет значение n_1, вторая n_2 и т. д., равна

$$P(n_1, n_2, \ldots, n_k) = \prod_{i=1}^{k} e^{-np_{0i}} (np_{0i})^{n_i}/n_i! = e^{-n} n^n \prod_{i=1}^{k} p_{0i}^{n_i}/n_i!. \tag{30.13}$$

Рассмотрим теперь условную вероятность наблюдения этих значений при фиксированной их сумме $\sum_{i=1}^{k} n_i = n$. Сумма k независимых пуассоновских случайных величин сама имеет распределение Пуассона с параметром $\sum_{i=1}^{k} np_{0i} = n$ (см. пример 11.11). Таким образом, вероятность того, что эта сумма равна n, есть

$$\mathsf{P}\left(\sum n_i = n\right) = e^{-n} n^n / n!. \tag{30.14}$$

Теперь мы можем получить требуемую условную вероятность:

$$P\left(n_1, n_2, \ldots, n_k \,\middle|\, \sum_i n_i = n\right) =$$

$$= \frac{P(n_1, n_2, \ldots, n_k)}{\mathsf{P}\left(\sum_i n_i = n\right)} = \frac{n!}{n_1!\, n_2! \ldots n_k!} p_{01}^{n_1} p_{02}^{n_2} \ldots p_{0k}^{n_k}. \tag{30.15}$$

Мы видим, что (30.15) есть в точности мультиномиальное распределение величин n_i, на которых основана наша процедура. Таким образом, в качестве альтернативы к выводу асимптотического распределения X^2, данному в примере 15.3 (см. **30.5**), мы можем получить это распределение, рассматривая n_i как значения k независимых пуассоновских случайных величин с параметрами np_{0i}, подчиненных условию $\sum n_i = n$. Согласно примеру 4.9 нормированная величина

$$x_i = \frac{n_i - np_{0i}}{(np_{0i})^{1/2}} \tag{30.16}$$

асимптотически нормальна при $n \to \infty$. Следовательно, при $n \to \infty$

$$X^2 = \sum_{i=1}^{k} x_i^2$$

представляет собой сумму квадратов k независимых нормальных величин, подчиненных единственному условию $\sum_i n_i = n$, что эквивалентно условию $\sum_i (np_{0i})^{1/2} x_i = 0$. Следовательно, согласно примеру 11.6 X^2 имеет асимптотически распределение χ^2 с $k - 1$ степенями свободы.

30.10 Польза этого второго доказательства состоит в том, что в совокупности с примером 11.6, на который мы ссылались, оно ясно показывает, что если наложить дополнительно s однородных линейных условий на n_i, то все влияние этого на асимптотическое распределение X^2 будет заключаться в уменьшении числа степеней свободы с $(k - 1)$ до $(k - s - 1)$.

Вернемся теперь к сложной гипотезе п. **30.8** в случае, когда выполняется (30.12). Предположим, что мы выбрали в качестве нашей совокупности $\boldsymbol{t}$ оценок параметра $\boldsymbol{\theta}$ оценки максимального правдоподобия (или другие эквивалентные асимптотически эффективные оценки), так что $\boldsymbol{t} = \hat{\boldsymbol{\theta}}$. В этом случае функция правдоподобия L есть просто мультиномиальное распределение (30.15), рассматриваемое как функция θ_j, от которых зависят p_{0i}. Таким образом,

$$\frac{\partial \log L}{\partial \theta_j} = \sum_{i=1}^{k} n_i \frac{\partial p_{0i}}{\partial \theta_j} \frac{1}{p_{0i}}, \qquad j = 1, 2, \ldots, s, \tag{30.17}$$

и МП-оценки в этом регулярном случае — это корни s уравнений, получаемых приравниванием (30.17) к нулю при каждом j. Ясно, что каждое такое уравнение есть однородное линейное соотношение между n_i. Мы, таким образом, видим, что в этом регулярном случае имеются s дополнительных ограничений, на-

лагаемых процессом эффективного оценивания параметра θ мультиномиального распределения, так что статистика (30.8) имеет асимптотически распределение χ^2 с $k-s-1$ степенями свободы. Более строгое и подробное доказательство дается в книге Крамера (1946), см. также Бёрч (1964b). Мы будем называть $\hat{\theta}$ мультиномиальной МП-оценкой.

Влияние оценивания на распределение X^2

30.11 Мы можем теперь, следуя Ватсону (1959), рассматривать общую задачу о влиянии оценивания неизвестных параметров на асимптотическое распределение статистики X^2. Мы ограничимся случаем, когда выполнено (30.12), и для любой оценки t параметра θ запишем *)

$$t - \theta = n^{-1/2} A x + o(n^{-1/2}), \tag{30.18}$$

где A — произвольная матрица $(s \times k)$ и x — вектор $(k \times 1)$ с i-м элементом

$$x_i = \frac{n_i - n p_{0i}(\theta)}{\{n p_{0i}(\theta)\}^{1/2}}, \tag{30.19}$$

определяемым так же, как в (30.16) для случая простой гипотезы; мы предполагаем, что A выбрана так, чтобы

$$\mathrm{M}(Ax) = 0. \tag{30.20}$$

Из (30.18) и (30.20) сразу следует, что матрица рассеяния t, $V(t)$, имеет порядок n^{-1}. Применяя разложение Тейлора к $\{p_{0i}(t) - p_{0i}(\theta)\}$ в (30.9), мы можем написать

$$X^2 = \sum_{i=1}^{k} y_i^2,$$

где при $n \to \infty$

$$y_i = x_i - n^{1/2} \sum_{j=1}^{s} (t_j - \theta_j) \frac{\partial p_{0i}(\theta)}{\partial \theta_j} \frac{1}{\{p_{0i}(\theta)\}^{1/2}} + o(1),$$

или, в матричной форме,

$$y = x - n^{1/2} B (t - \theta) + o(1), \tag{30.21}$$

где B — матрица $(k \times s)$ с (i, j)-м элементом

$$b_{ij} = \frac{\partial p_{0i}(\theta)}{\partial \theta_j} \frac{1}{\{p_{0i}(\theta)\}^{1/2}}. \tag{30.22}$$

*) Соотношение (30.18) возможно лишь для оценок, выражаемых через частоты $n_1, \ldots, n_k$ (см. (30.19)). Оценка t, зависящая от индивидуальных значений наблюдений, как, например, обычная МП-оценка (см. **30.17**), может не допускать такого представления. (*Прим. ред.*).

Подставляя (30.18) в (30.21), мы находим, что

$$\boldsymbol{y} = (\boldsymbol{I} - \boldsymbol{BA})\,\boldsymbol{x} + o(1). \tag{30.23}$$

30.12 В силу уравнения (30.19) x_i имеют нулевые средние. Согласно многомерной центральной предельной теореме при $n \to \infty$ они имеют в пределе многомерное нормальное распределение, и их матрица рассеяния равна, если обозначить временно $p_{0i}(\theta)$ через p_i,

$$\boldsymbol{V}(\boldsymbol{x}) = \begin{pmatrix} 1-p_1 & -(p_1p_2)^{1/2} & -(p_1p_3)^{1/2} & \dots & -(p_1p_k)^{1/2} \\ -(p_1p_2)^{1/2} & 1-p_2 & -(p_2p_3)^{1/2} & \dots & -(p_2p_k)^{1/2} \\ \dots & \dots & \dots & \dots & \dots \\ -(p_kp_1)^{1/2} & -(p_kp_2)^{1/2} & -(p_kp_3)^{1/2} & \dots & 1-p_k \end{pmatrix} =$$

$$= \boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})', \tag{30.24}$$

где $\boldsymbol{p}^{1/2}$ — вектор $(k \times 1)$ с i-м элементом $\{p_{0i}(\theta)\}^{1/2}$. Из (30.23) и (30.24) сразу следует, что y_i также асимптотически нормальны с нулевыми средними и матрицей рассеяния

$$\boldsymbol{V}(\boldsymbol{y}) = (\boldsymbol{I} - \boldsymbol{BA})\{\boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'\}(\boldsymbol{I} - \boldsymbol{BA})'. \tag{30.25}$$

Таким образом, $X^2 = \boldsymbol{y}'\boldsymbol{y}$ распределено асимптотически как сумма квадратов k нормальных случайных величин с нулевыми средними и матрицей рассеяния (30.25). Как мы знаем из **15.10—11**, X^2 имеет распределение χ^2 с r степенями свободы тогда и только тогда, когда $\boldsymbol{V}(y)$ идемпотентна и r характеристических чисел ее равны единице, а $(k - r)$ — нулю, так что ее след равен r.

30.13 Теперь мы рассмотрим частные случаи (30.25). Прежде всего, случай простой гипотезы, где не требуется никакой оценки, можно формально получить, полагая $\boldsymbol{A} \equiv \boldsymbol{0}$ в (30.18). Тогда (30.25) становится равной просто

$$\boldsymbol{V}(\boldsymbol{y}) = \boldsymbol{V}(\boldsymbol{x}) = \boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'. \tag{30.26}$$

Так как $(\boldsymbol{p}^{1/2})'(\boldsymbol{p}^{1/2}) = \sum_{i=1}^{k} p_{0i}(\theta) = 1$, то, возводя матрицу (30.26) в квадрат, мы видим, что она идемпотентна и ее след равен $(k - 1)$. Таким образом, X^2 в этом случае имеет распределение χ^2_{k-1}, как мы уже знаем из двух других доказательств.

30.14 Случай сложной гипотезы уже не столь прост. Предположим сначала, как в **30.10**, что используются мультиномиальные МП-оценки $\hat{\theta}$. Мы ищем вид матрицы $\boldsymbol{A}$ в (30.8), когда $t = \hat{\theta}$. Из **18.26** мы знаем, что матрица, обратная к матрице рассеяния $\hat{\theta}$, имеет асимптотически элементы

$$\{\boldsymbol{V}(\hat{\theta})\}^{-1}_{jl} = -\mathbf{M}\left\{\frac{\partial^2 \log L}{\partial\theta_j\,\partial\theta_l}\right\}, \quad j, l = 1, 2, \dots, s. \tag{30.27}$$

Согласно (30.17) мультиномиальные уравнения МП дают

$$\frac{\partial^2 \log L}{\partial\theta_j\,\partial\theta_l} = \sum_{i=1}^{k} \frac{n_i}{p_{0i}} \left(\frac{\partial^2 p_{0i}}{\partial\theta_j\,\partial\theta_l} - \frac{1}{p_{0i}} \frac{\partial p_{0i}}{\partial\theta_j} \frac{\partial p_{0i}}{\partial\theta_l} \right). \tag{30.28}$$

Беря математические ожидания в (30.28), мы находим

$$-\mathbf{M}\left\{ \frac{\partial^2 \log L}{\partial\theta_j\,\partial\theta_l} \right\} = n \left\{ \sum_{i=1}^{k} \frac{1}{p_{0i}} \frac{\partial p_{0i}}{\partial\theta_j} \frac{\partial p_{0i}}{\partial\theta_l} - \sum_{i=1}^{k} \frac{\partial^2 p_{0i}}{\partial\theta_j\,\partial\theta_l} \right\}. \tag{30.29}$$

Второй член в правой части (30.29) обращается в нуль, так как он равен $\dfrac{\partial^2}{\partial\theta_j\,\partial\theta_l} \sum_{i=1}^{k} p_{0i}$. Таким образом, пользуясь (30.22), имеем

$$-\mathbf{M}\left\{ \frac{\partial^2 \log L}{\partial\theta_j\,\partial\theta_l} \right\} = n \sum_{i=1}^{k} b_{ij} b_{il},$$

так что из (30.27)

$$\{\boldsymbol{V}(\hat{\theta})\}^{-1} = n\boldsymbol{B}'\boldsymbol{B},$$

или

$$\boldsymbol{C} = n\boldsymbol{V}(\hat{\theta}) = (\boldsymbol{B}'\boldsymbol{B})^{-1}. \tag{30.30}$$

Но из (30.18) и (30.24) мы имеем

$$\boldsymbol{D} = n\boldsymbol{V}(\boldsymbol{t}) = \boldsymbol{A}\boldsymbol{V}(\boldsymbol{x})\boldsymbol{A}' = \boldsymbol{A}\{\boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'\}\boldsymbol{A}'. \tag{30.31}$$

Здесь (30.30) и (30.31) — два различных выражения для одной и той же матрицы.

Чтобы найти $\boldsymbol{A}$ так, чтобы выполнялось (30.30), заметим, что

$$\boldsymbol{B}'(\boldsymbol{p}^{1/2}) = \boldsymbol{0} \tag{30.32}$$

$\left(\text{так как } j\text{-й элемент этого } (s+1)\text{-вектора равен } \dfrac{\partial}{\partial\theta_j} \sum_{i=1}^{k} p_{0i} = 0\right)$,

и, следовательно, если $\boldsymbol{A} = \boldsymbol{G}\boldsymbol{B}'$, где $\boldsymbol{G}$ симметрична и невырождена, то (30.31) дает $\boldsymbol{G}\boldsymbol{B}'\boldsymbol{B}\boldsymbol{G}'$. Если это должно быть равно (30.30), то мы имеем, очевидно, $\boldsymbol{G} = (\boldsymbol{B}'\boldsymbol{B})^{-1}$, и окончательно

$$\boldsymbol{A} = (\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}' \tag{30.33}$$

в случае мультиномиальной МП-оценки. Тогда, пользуясь (30.32), приводим (30.25) к виду

$$\boldsymbol{V}(\boldsymbol{y}) = \{\boldsymbol{I} - \boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}'\}^2 - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})' =$$
$$= \boldsymbol{I} - \boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^{-1}\boldsymbol{B}' - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'. \tag{30.34}$$

Возведением в квадрат можно показать, что эта матрица идемпотентна. Ее ранг равен ее следу, который, как в **19.9**, дается

формулой

$$\operatorname{tr} \boldsymbol{V}(\boldsymbol{y}) = \operatorname{tr}\{\boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'\} - \operatorname{tr} \boldsymbol{B}' \cdot \boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^{-1},$$

и в силу **30.13** это равно

$$\operatorname{tr} \boldsymbol{V}(\boldsymbol{y}) = (k-1) - s.$$

Таким образом, X^2 имеет асимптотически распределение χ^2_{k-s-1}, как мы видели в **30.10**.

30.15 Применяемый сейчас подход дает нам возможность дальнейшего исследования: что случится с асимптотическим распределением статистики X^2, если применять какие-нибудь оценки, отличные от $\hat{\theta}$? Этот вопрос был впервые рассмотрен в простейшем случае Фишером (1928b). Чернов и Леман (1954) рассмотрели случай, представляющий особый интерес, когда применяются МП-оценки, основанные на n индивидуальных наблюдениях, а не мультиномиальные МП-оценки $\hat{\theta}$, основанные на k частотах n_i, которые мы рассматривали до сих пор. Если мы имеем значения n наблюдений, то ясно, что использование этого знания при оценке θ дает эффективную процедуру *), даже если мы в критерии согласия собираемся пользоваться лишь частотами k классов. Мы, однако, увидим, что статистика X^2, полученная таким образом, уже не имеет асимптотически распределения χ^2.

30.16 Вернемся к общему выражению (30.25) для матрицы рассеяния. Произведя умножение, мы перепишем его в виде

$$\boldsymbol{V}(\boldsymbol{y}) = \{\boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'\} - \boldsymbol{B}\boldsymbol{A}\{\boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'\} - \\ - \{\boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'\}\boldsymbol{A}'\boldsymbol{B}' + \boldsymbol{B}\boldsymbol{A}\{\boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'\}\boldsymbol{A}'\boldsymbol{B}'. \quad (30.35)$$

Вместо того, чтобы находить характеристические числа λ_i матрицы $\boldsymbol{V}(\boldsymbol{y})$, мы рассмотрим характеристические числа матрицы $\boldsymbol{I} - \boldsymbol{V}(\boldsymbol{y})$, равные $1 - \lambda_i$. Мы запишем эту матрицу в виде

$$\boldsymbol{I} - \boldsymbol{V}(\boldsymbol{y}) = (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})' + \\ + \boldsymbol{B}\left[\boldsymbol{A}\{\boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'\} - \tfrac{1}{2}\boldsymbol{A}\{\boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'\}\boldsymbol{A}'\boldsymbol{B}'\right] + \\ + \left[\{\boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'\}\boldsymbol{A}' - \tfrac{1}{2}\boldsymbol{B}\boldsymbol{A}\{\boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'\}\boldsymbol{A}'\right]\boldsymbol{B}' = \\ = (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})' + \boldsymbol{B}\left[\boldsymbol{A}\{\boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'\} - \tfrac{1}{2}\boldsymbol{D}\boldsymbol{B}'\right] + \\ + \left[\{\boldsymbol{I} - (\boldsymbol{p}^{1/2})(\boldsymbol{p}^{1/2})'\}\boldsymbol{A}' - \tfrac{1}{2}\boldsymbol{B}\boldsymbol{D}\right]\boldsymbol{B}'. \quad (30.36)$$

*) Это интуитивное представление оказывается неверным. Критерий, использующий МП-оценку по n наблюдениям, при различных альтернативах может иметь как бо́льшую, так и меньшую мощность по сравнению с критерием, использующим оценку $\hat{\theta}$, **см.** Д. М. Чибисов (1971), Теория вероят. и ее примен. **XVI**, 3—20. (*Прим. ред.*)

После подстановки (30.31) (30.36) может быть переписано в виде произведения двух блочных матриц:

$$\boldsymbol{I}-\boldsymbol{V}(\boldsymbol{y})=\begin{Vmatrix} (\boldsymbol{p}^{1/2}) \\ \boldsymbol{B} \\ \{\boldsymbol{I}-(\boldsymbol{p}^{1/2})\,(\boldsymbol{p}^{1/2})'\}\,\boldsymbol{A}'-\frac{1}{2}\,\boldsymbol{BD} \end{Vmatrix} \begin{Vmatrix} (\boldsymbol{p}^{1/2})' \\ \boldsymbol{A}\{\boldsymbol{I}-(\boldsymbol{p}^{1/2})\,(\boldsymbol{p}^{1/2})'\}-\frac{1}{2}\,\boldsymbol{DB}' \\ \boldsymbol{B}' \end{Vmatrix}'.$$

Матрицы в правой части можно переставить, не изменяя ненулевых характеристических чисел. При этом их произведение из матрицы $(k\times k)$ становится матрицей $(2s+1)\times(2s+1)$, которую, пользуясь (30.30) и (30.32), приводим к виду

$$\left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{$\boldsymbol{M}$}} \\ 0 & & \end{array}\right)=$$

$$=\left\|\begin{array}{c|cc} 1 & 0_{1\times s} & 0_{1\times s} \\ \hline 0_{s\times 1} & \boldsymbol{AB}-\frac{1}{2}\,\boldsymbol{DC}^{-1} & \boldsymbol{D}-\frac{1}{2}\,\boldsymbol{ABD}-\frac{1}{2}\,\boldsymbol{DB}'\boldsymbol{A}+\frac{1}{4}\,\boldsymbol{DC}^{-1}\boldsymbol{D} \\ 0_{s\times 1} & \boldsymbol{C}^{-1} & \boldsymbol{B}'\boldsymbol{A}'-\frac{1}{2}\,\boldsymbol{C}^{-1}\boldsymbol{D} \end{array}\right\|. \quad (30.37)$$

Одно из характеристических чисел матрицы (30.37) равно единице, а $2s$ остальных являются характеристическими числами матрицы $\boldsymbol{M}$, выделенной в ее правом нижнем углу. Если $k\geqslant 2s+1$, как почти всегда бывает в приложениях, то отсюда следует, что (30.36) имеет $k-2s-1$ нулевых характеристических чисел, одно, равное единице, а $2s$ остальных являются характеристическими числами $\boldsymbol{M}$. Таким образом, для самой матрицы $\boldsymbol{V}(\boldsymbol{y})$ мы имеем $k-2s-1$ характеристических чисел, равных единице, одно нулевое и $2s$, являющихся дополнениями до единицы характеристических чисел матрицы $\boldsymbol{M}$.

30.17 Мы рассмотрим теперь задачу, введенную в **30.15**. Предположим, что для оценки θ мы пользуемся МП-оценкой, основанной на n индивидуальных наблюдениях, которую мы будем называть «обычной МП-оценкой» и обозначать $\hat{\theta}^*$. Мы знаем из **18.26**, что если f — функция плотности наблюдений, то мы имеем асимптотически

$$\boldsymbol{D}=n\boldsymbol{V}(\hat{\theta}^*)=-\left\{\mathsf{M}\left(\frac{\partial^2\log f}{\partial\theta_j\,\partial\theta_l}\right)\right\}^{-1} \quad (30.38)$$

и элементы $\hat{\theta}^*$ являются корнями уравнения

$$\frac{\partial\log L}{\partial\theta_i}=0, \qquad i=1,\ 2,\ \ldots,\ s,$$

где L теперь — обычная (не мультиномиальная) функция правдоподобия. Таким образом, если θ_0 — истинное значение, то мы имеем разложение Тейлора

$$0 = \left[\frac{\partial \log L}{\partial \theta_i}\right]_{\theta=\hat{\theta}^*} =$$

$$= \left[\frac{\partial \log L}{\partial \theta_i}\right]_{\theta=\theta_0} + (\hat{\theta}_j^* - \theta_{0j}) \left[\frac{\partial^2 \log L}{\partial \theta_i \, \partial \theta_j}\right]_{\theta=\theta_0+\varepsilon}, \qquad i, j = 1, 2, \ldots, s,$$

и, как в **18.26**, получаем, пользуясь (30.38), что асимптотически

$$\hat{\theta}^* - \theta = \frac{1}{n} \boldsymbol{D} \begin{pmatrix} \frac{\partial \log L}{\partial \theta_1} \\ \vdots \\ \frac{\partial \log L}{\partial \theta_s} \end{pmatrix}, \tag{30.39}$$

где θ — вектор истинных значений. Поскольку обе совокупности МП-оценок состоятельны, $\frac{\partial \log L}{\partial \theta_j}$ будет в первом приближении эквивалентно выражению (30.17) *), которое вследствие (30.22) равно

$$\sum_{i=1}^{k} \frac{n_i}{(p_{0i})^{1/2}} b_{ij} = n^{1/2} \sum_{i=1}^{k} \frac{n_i}{(np_{0i})^{1/2}} b_{ij},$$

и мы можем записать это в обозначениях (30.19) как

$$n^{1/2} \sum_{i=1}^{k} x_i b_{ij} = n^{1/2} (\boldsymbol{B}'\boldsymbol{x})_j.$$

Таким образом, (30.39) принимает вид

$$\hat{\theta}^* = \theta + n^{-1/2} \boldsymbol{D}\boldsymbol{B}'\boldsymbol{x} \tag{30.40}$$

и сравнение (30.40) с (30.18) показывает, что здесь

$$\boldsymbol{A} = \boldsymbol{D}\boldsymbol{B}'. \tag{30.41}$$

---

*) Это утверждение и вытекающие из него формулы (30.40) и (30.41) неверны, так, из (30.40) в силу (30.24) и (30.32) следует

$$n\boldsymbol{V}(\hat{\theta}^*) = \boldsymbol{D}\boldsymbol{B}'\boldsymbol{V}(\boldsymbol{x})\,\boldsymbol{B}\boldsymbol{D}' = \boldsymbol{D}\boldsymbol{B}'\boldsymbol{B}\boldsymbol{D}',$$

что не совпадает с (30.38). Ошибочность рассуждения здесь связана с невозможностью представления (30.18) для оценки $\hat{\theta}^*$ (см. примечание на стр. 567). Формула (30.42) и дальнейшие выводы в **30.18** тем не менее верны; доказательства см. в работах Чернова и Лемана (1954) и Чибисова (цит. на стр. 570). (*Прим. ред.*)

30.18 Матрица рассеяния (30.25) теперь принимает вид

$$\boldsymbol{V}(y) = (\boldsymbol{I} - \boldsymbol{BDB}')\{\boldsymbol{I} - (p^{1/2})(p^{1/2})'\}(\boldsymbol{I} - \boldsymbol{BDB}'),$$

откуда, пользуясь (30.31), (30.32) и (30.41), получаем

$$\boldsymbol{V}(y) = \boldsymbol{I} - (p^{1/2})(p^{1/2})' - \boldsymbol{BDB}'. \qquad (30.42)$$

Матрица (30.42) не идемпотентна, что проверяется возведением ее в квадрат. Кроме того, все три матрицы в правой части неотрицательно определены, и мы можем записать

$$\boldsymbol{D} + \boldsymbol{P} = \boldsymbol{C},$$

где $\boldsymbol{P}$ неотрицательно определена, так как $\boldsymbol{D}$ — матрица рассеяния асимптотически эффективных обычных МП-оценок, а $\boldsymbol{C}$ — матрица рассеяния мультиномиальных МП-оценок, каждый из диагональных элементов которой не может быть меньше соответствующего элемента $\boldsymbol{D}$. Таким образом, (30.42) может быть записано как

$$\boldsymbol{V}(y) = \boldsymbol{I} - (p^{1/2})(p^{1/2})' - \boldsymbol{BCB}' + \boldsymbol{BPB}'. \qquad (30.43)$$

Первые два члена в правой части (30.43) — это то, что мы получили в (30.26) для случая, где не производится никакой оценки, когда $k - 1$ характеристических чисел $\boldsymbol{V}(y)$ равны единице и одно равно нулю. Первые три члена дают (30.34), где при применении мультиномиальных МП-оценок $\boldsymbol{V}(y)$ имеет $k-s-1$ характеристических чисел, равных единице, и $s+1$ нулевых. Вследствие неотрицательной определенности всех членов приведение (30.43) к каноническому виду показывает, что характеристические числа (30.43) ограничены соответствующими характеристическими числами (30.26) и (30.34). Таким образом, $k - s - 1$ характеристических чисел матрицы (30.43) равны единице, одно нулю и s заключены между нулем и единицей, как установлено Черновым и Леманом (1954).

Из того факта, что при возрастании k две совокупности МП-оценок $\hat{\theta}$ и $\hat{\theta}^*$ сближаются между собой, так что $\boldsymbol{D} \to \boldsymbol{C}$, следует, что последние s характеристических чисел стремятся к нулю при $k \to \infty$.

30.19 Итак, результат, который мы получили, состоит в том, что X^2 не имеет асимптотически χ^2-распределения, когда для оценки параметров используются вполне эффективные оценки (обычного максимального правдоподобия). Однако распределение X^2 заключено между χ^2_{k-1} и χ^2_{k-s-1}, и при возрастании k они становятся столь близки между собой, что различием можно пренебречь, — это другое выражение для заключительного утверждения в **30.18**. Но при малых k использование распределения χ^2_{k-s-1} для проверки гипотез может привести к серьезной

ошибке, так как вероятность превышения любого данного значения будет больше, чем мы предполагаем. s редко бывает больше, чем 1 или 2, но при использовании обычной МП-оценки полезно убедиться, что X^2 превышает оба критических значения для χ^2_{k-s-1} и χ^2_{k-1}. Таблицы распределения χ^2 показывают, что для критерия с уровнем значимости $\alpha = 0{,}05$ критическое значение при $k-1$ степенях свободы превышает критическое значение при $k-s-1$ степенях свободы, если s мало, приблизительно на Cs, где C убывает от примерно 1,5 при $(k-s-1) = 5$ до примерно 1,2 при $(k-s-1) = 30$. При $\alpha = 0{,}01$ соответствующие значения C равны приближенно 1,7 и 1,3.

Выбор классов для критерия X^2

30.20 Вся асимптотическая теория критерия X^2, которая обсуждалась до сих пор, справедлива при любом определении k классов, в которые группируются наблюдения, если только они определяются *безотносительно к результатам наблюдений.* Выделенное условие существенно, так как мы не предусматривали возможности того, чтобы границы классов сами были случайными величинами. Однако обычной практикой является определение границ классов, а иногда даже выбор k по рассмотрении общей картины, даваемой наблюдениями. Мы должны, следовательно, обсудить способы образования классов и затем рассмотреть, насколько это влияет на развитую нами теорию.

Мы рассмотрим сначала определение границ классов, оставляя вопрос о выборе k на более позднее время. Если сама задача обусловливает дискретный характер данных (как в примере 30.1, где было четыре естественных группы) или если мы имеем выборку наблюдений из дискретного распределения, то задача о границах классов возникает только в том смысле, что мы можем решить (чтобы уменьшить k или чтобы улучшить точность асимптотического распределения X^2, как мы увидим ниже в **30.30**) объединить некоторые из теоретических частот в дискретных точках. Действительно, если дискретное распределение имеет бесконечную область, как, например, пуассоновское, мы вынуждены, если k не должно быть бесконечным, производить объединение некоторых классов с чрезвычайно малыми, в большинстве, теоретическими частотами. Но проблема границ классов наиболее остро возникает, только когда мы имеем выборку из непрерывного распределения. В этом случае нет никаких теоретических частот. Если мы предположим, что k каким-то образом определено заранее, то как определять границы классов?

На практике решение обычно диктуется арифметическим удобством: классы берутся покрывающими равные отрезки из-

менения варианты, за исключением крайних, где область изменения варианты бесконечна. Размер класса грубо определяется дисперсией распределения, тогда как положение распределения помогает определить, где должен быть центральный класс. Таким образом, если нам нужно было бы образовать 10 классов для выборки, проверяемой на нормальность, мы могли бы грубо оценить (возможно, на глаз) выборочные среднее $\bar{x}$ и стандартное отклонение s и взять в качестве границ классов $\bar{x} \pm \frac{1}{2}sj$, $j = 1, 2, 3, 4$. Тогда получились бы классы

$$(-\infty, \bar{x}-2s), \quad (\bar{x}-2s, \bar{x}-1{,}5s), \quad (\bar{x}-1{,}5s, \bar{x}-s),$$

$$(\bar{x}-s, \bar{x}-0{,}5s), \quad (\bar{x}-0{,}5s, \bar{x}),$$

$$(\bar{x}, \bar{x}+0{,}5s), \quad (\bar{x}+0{,}5s, \bar{x}+s), \quad (\bar{x}+s, x+1{,}5s),$$

$$(\bar{x}+1{,}5s, \bar{x}+2s), \quad (\bar{x}+2s, \infty).$$

30.21 Хотя это и не очень точно определенная процедура, ясно, что она делает границы классов случайными величинами, и заранее не очевидно, что статистика X^2, вычисленная для классов, образованных таким способом, имеет то же асимптотическое распределение, что и в случае, когда классы фиксированы наперед. Однако интуиция подсказывает, что, поскольку асимптотическая теория справедлива для любой совокупности из k фиксированных классов, она должна быть справедлива и когда границы классов определяются по выборке. То, что это утверждение выполняется при определении границ классов регулярными оценками неизвестных параметров, было показано для случая нормального распределения Ватсоном (1957b) и для общего случая непрерывных распределений Роем (1956) и Ватсоном (1958, 1959) *).

Таким образом, мы можем пренебречь случайностью границ интервалов, коль скоро дело касается асимптотического распределения X^2 при гипотезе H_0. Конечно, это повлияет на распределения при малых выборках, но о точных размерах этого влияния еще ничего не известно. (Распределения X^2 при малых выборках в случае фиксированных границ мы обсуждаем ниже, в **30.30**).

Метод равных вероятностей для построения классов

30.22 Мы можем теперь непосредственно обратиться к вопросу о том, как следует определять границы классов в свете утверждения последнего абзаца в **30.21**. Искать оптимальный

*) См. также работу Чибисова, цит. в примечании на стр. 570. (*Прим. ред.*)

метод определения границ мы должны в терминах мощности критерия; надо выбрать такую совокупность границ, которая максимизировала бы мощность для критерия данного размера. К сожалению, метода, пригодного для этого, пока нет, хотя можно надеяться, что недавнее возобновление интереса к теории критериев X^2 будет стимулировать исследования в этой области. Мы должны, следовательно, искать какие-то средства избежать того неприятного факта, что имеется многообразие возможных совокупностей классов, каждая из которых будет давать, вообще говоря, иной результат для тех же данных; нам нужно правило, приемлемое и применимое для практики.

Одно такое правило было предложено Манном и Вальдом (1942) и Гамбелом (1943): для данного k выбрать классы так, чтобы все гипотетические вероятности p_{0i} были равны $1/k$. Это совершенно определенная и однозначная процедура. Она отличается арифметически от обычного метода, описанного в **30.20** (в котором классами служат интервалы изменения варианты равной длины), тем, что нам приходится пользоваться таблицами, чтобы обеспечить равенство p_{0i}. Для точного осуществления этого нужно, чтобы имеющиеся данные были не группированы. Эта процедура иллюстрируется примером 30.2.

Пример 30.2

Кэнуй (1959) приводит, с точностью до сдвига, 1000 случайных отклонений, подчиняющихся распределению

$$dF = \exp(-x)\,dx, \qquad 0 \leqslant x \leqslant \infty.$$

Первые 50 из них, упорядоченные по величине, равны: 0,01; 0,01; 0,04; 0,17; 0,18; 0,22; 0,22; 0,25; 0,25; 0,29; 0,42; 0,46; 0,47; 0,47; 0,56; 0,59; 0,67; 0,68; 0,70; 0,72; 0,76; 0,78; 0,83; 0,85; 0,87; 0,93; 1,00; 1,01; 1,01; 1,02; 1,03; 1,05; 1,32; 1,34; 1,37; 1,47; 1,50; 1,52; 1,54; 1,59; 1,71; 1,90; 2,10; 2,35; 2,46; 2,46; 2,50; 3,73; 4,07; 6,03.

Допустим, что мы хотим образовать четыре класса для критерия X^2. Естественной группировкой с интервалами равной длины была бы:

Значения варианты	Наблюденная частота	Гипотетическая частота
0—0,50	14	19,7
0,51—1,00	13	11,9
1,01—1,50	10	7,2
1,51 и более	13	11,2
	50	50,0

Гипотетические частоты получены из таблиц в *Biometrika Tables* функции распределения величины χ^2 с 2 степенями свободы, представляющий собой удвоенную случайную величину с указанным выше распределением. Мы находим $X^2 = 3{,}1$ с 3 степенями свободы, значение, при котором гипотетическое распределение не отвергалось бы никаким критерием размера меньшего, чем $\alpha = 0{,}37$; согласие между наблюдениями и гипотезой, следовательно, весьма удовлетворительное.

Рассмотрим теперь, как те же данные обрабатывались бы по методу из **30.22**. Сначала мы определяем значения гипотетической величины, разбивающие ее на четыре равновероятных класса, — это, разумеется, квартили. Таблицы *Biometrika Tables* дают значения 0,228, 0,693, 1,386. Теперь мы составляем таблицу:

Значения варианты	Наблюденная частота	Гипотетическая частота
0—0,28	9	12,5
0,29—0,69	9	12,5
0,70—1,38	17	12,5
1,39 и более	15	12,5
	50	50,0

Статистику X^2 теперь легче вычислять, так как (30.8) сводится к

$$X^2 = \frac{k}{n} \sum_{i=1}^{k} n_i^2 - n, \tag{30.44}$$

поскольку все гипотетические вероятности $p_{0i} = 1/k$. Мы находим, что $X^2 = 3{,}9$; при таком значении X^2 гипотеза не будет отвергнута, если размер критерия не превышает 0,27. Это по-прежнему весьма удовлетворительный результат, но критерий с равными вероятностями кажется более критичным к гипотезе, чем был первый критерий.

В этом примере не было параметров, подлежащих оценке. Если нужно оценивать s параметров, мы неизбежно сталкиваемся с проблемой, рассмотренной в **30.15—19**: при использовании обычных МП-оценок применимы выводы п. **30.19**.

30.23 Метод образования равновероятных классов для X^2-критерия, обладая тем достоинством, что он устраняет неопределенность в решении о границах классов, не обязательно увеличивает мощность критерия, так как можно ожидать, что

гипотеза согласия наиболее уязвима на краях области изменения варианты, и метод равных вероятностей вполне может привести к потере чувствительности на краях, если только k не будет достаточно велико. Это подводит нас к вопросу о том, как следует выбирать k, и, чтобы обсуждать этот вопрос, мы должны рассмотреть мощность X^2-критерия. Прежде всего мы исследуем моменты статистики X^2.

Моменты статистики критерия X^2

30.24 Мы предполагаем, как прежде, что при гипотезе H_0 имеются гипотетические вероятности p_{0i}, так что статистикой является, как в (30.8),

$$X^2 = \frac{1}{n}\sum_{i=1}^{k}\frac{n_i^2}{p_{0i}} - n.$$

Мы ограничимся простой гипотезой. Предположим теперь, что истинные вероятности равны p_{1i}, $i = 1, 2, \ldots, k$. Тогда математическое ожидание статистики критерия равно

$$\mathsf{M}(X^2) = \frac{1}{n}\sum_{i=1}^{k}\frac{1}{p_{0i}}\mathsf{M}(n_i^2) - n.$$

По формуле (5.80) для моментов мультиномиального распределения

$$\mathsf{M}(n_i^2) = np_{1i}(1 - p_{1i}) + n^2p_{1i}^2, \tag{30.45}$$

откуда

$$\mathsf{M}(X^2) = \sum_{i=1}^{k}\frac{p_{1i}(1-p_{1i})}{p_{0i}} + n\left\{\sum_{i=1}^{k}\frac{p_{1i}^2}{p_{0i}} - 1\right\}. \tag{30.46}$$

Когда справедлива гипотеза H_0, это сводится к

$$\mathsf{M}(X^2 \mid H_0) = k - 1. \tag{30.47}$$

Это точный результат, об асимптотической справедливости которого нам уже известно, так как X^2 тогда распределена как χ^2_{k-1}. Если мы продифференцируем (30.46) по p_{1i}, подчиненным условию $\sum_i p_{1i} = 1$, то мы найдем, что при $n \to \infty$ (30.46) достигает минимума, когда $p_{1i} = p_{0i}$. Для любой гипотезы H_1, задающей набор вероятностей $p_{1i} \neq p_{0i}$, мы, следовательно, имеем асимптотически

$$\mathsf{M}(X^2 \mid H_1) > k - 1. \tag{30.48}$$

Как и асимптотические рассуждения, основанные на статистике ОП в **30.6**, (30.48) указывает, что критическая область для

X^2-критерия состоит из больших значений, хотя это и не является окончательным аргументом, так как X^2 при гипотезе H_1 уже не имеет асимптотического распределения χ^2. На самом деле этим распределением будет при условиях, даваемых далее в **30.27**, нецентральное распределение χ^2.

Даже дисперсия X^2 сравнительно сложно зависит от p_{0i} и p_{1i} (см. упражнение 30.5). Однако в случае равных вероятностей ($p_{0i} = 1/k$) асимптотическая дисперсия заметно упрощается, и мы находим (доказательство оставляем читателю в качестве упражнения 30.3)

$$\left.\begin{aligned} &\mathsf{D}(X^2 \mid H_0) = 2(k-1), \\ &\mathsf{D}(X^2 \mid H_1) = 2k^2\left\{\sum_i p_{1i}^2 + 2(n-1)\sum_i p_{1i}^3 - (2n-1)\left(\sum_i p_{1i}^2\right)^2\right\}. \end{aligned}\right\} \quad (30.49)$$

Из (30.46) мы имеем также в случае равных вероятностей

$$\mathsf{M}(X^2 \mid H_1) = (k-1) + (n-1)\left(k\sum_i p_{1i}^2 - 1\right). \quad (30.50)$$

Выражение (30.50) больше, чем $(k-1)$, при любом n.

Состоятельность и несмещенность критерия X^2

30.25 Равенства (30.49) и (30.50) достаточны, чтобы показать состоятельность X^2-критерия с равными вероятностями. Действительно, критерий состоит в сравнении значения X^2 с фиксированным критическим значением, скажем c_α, на верхнем «хвосте» своего распределения. Когда справедлива гипотеза H_1, среднее и дисперсия X^2 имеют порядок n. По неравенству Чебышева (3.95)

$$\mathsf{P}\{|X^2 - \mathsf{M}(X^2)| \geqslant \lambda[\mathsf{D}(X^2)]^{1/2}\} \leqslant 1/\lambda^2. \quad (30.51)$$

Поскольку c_α фиксировано, оно отличается от $\mathsf{M}(X^2)$ на величину порядка n, так что если нас интересует вероятность того, что X^2 отличается от своего среднего настолько, чтобы попасть ниже c_α, то множитель λ в левой части (30.51) должен быть порядка $n^{1/2}$, и, следовательно, правая часть имеет порядок n^{-1}. Таким образом,

$$\lim_{n\to\infty} \mathsf{P}\{X^2 < c_\alpha\} = 0,$$

и критерий состоятелен против любой альтернативы H_1, задающей неравные вероятности классов.

Общий X^2-критерий с неравными p_{0i} также состоятелен против альтернатив, задающих хотя бы одно $p_{1i} \neq p_{0i}$, что понятно интуитивно. Доказательство было дано Нейманом (1949).

30.26 Хотя X^2-критерий состоятелен (и, следовательно, асимптотически несмещен), нет оснований ожидать, что он будет несмещенным в общем случае против близких альтернатив при малых n. Однако Манн и Вальд (1942) показали, что критерий с равными вероятностями является локально несмещенным. Обозначим P мощность критерия и разложим P в ряд Тейлора с k значениями $\theta_i = p_i - (1/k)$ в качестве аргументов. Мы имеем тогда

$$P(\theta_1, \theta_2, \ldots, \theta_k) = P(0, 0, \ldots, 0) + \sum \theta_i \frac{\partial P}{\partial \theta_i} + \\ + \frac{1}{2}\left\{\sum_i \theta_i^2 \frac{\partial^2 P}{\partial \theta_i^2} + \sum_{i \neq j}\sum \theta_i \theta_j \frac{\partial^2 P}{\partial \theta_i \partial \theta_j}\right\}, \quad (30.52)$$

где все производные берутся в точке $(0, 0, \ldots, 0)$, отвечающей гипотезе H_0. Для критерия размера α

$$P(0, 0, \ldots, 0) = \alpha.$$

Далее, поскольку P — симметричная функция аргументов θ_i, все $\frac{\partial P}{\partial \theta_i}$ равны между собой в точке $(0, 0, \ldots, 0)$, и то же верно для $\frac{\partial^2 P}{\partial \theta_i^2}$ и $\frac{\partial^2 P}{\partial \theta_i \partial \theta_j}$. Мы можем, следовательно, переписать (30.52) в виде

$$P = \alpha + \frac{\partial P}{\partial \theta_1}\sum_i \theta_i + \frac{1}{2}\left\{\frac{\partial^2 P}{\partial \theta_1^2}\sum_i \theta_i^2 + \frac{\partial^2 P}{\partial \theta_1 \partial \theta_2}\sum_{i \neq j}\sum \theta_i \theta_j\right\}. \quad (30.53)$$

Далее,

$$0 = \sum_i \theta_i = \left(\sum_i \theta_i\right)^2 = \sum_i \theta_i^2 + \sum_{i \neq j}\sum \theta_i \theta_j.$$

Поэтому выражение (30.53) равно

$$P = \alpha + \frac{1}{2}\left(\frac{\partial^2 P}{\partial \theta_1^2} - \frac{\partial^2 P}{\partial \theta_1 \partial \theta_2}\right)\sum_i \theta_i^2 + \ldots \quad (30.54)$$

Мы можем найти значения производных второго порядка в (30.54) непосредственно из точного выражения для мощности

$$P = \sum_{X^2 \geqslant c_\alpha} \frac{n!}{n_1!\, n_2! \ldots n_k!} p_1^{n_1} p_2^{n_2} \ldots p_k^{n_k}. \quad (30.55)$$

Поскольку $\theta_i = p_i - \frac{1}{k}$, мы находим из (30.55)

$$\frac{\partial^2 P}{\partial \theta_1^2} - \frac{\partial^2 P}{\partial \theta_1 \partial \theta_2} = \sum \{n_1(n_1 - 1) - n_1 n_2\} \frac{n!}{n_1!\, n_2! \ldots n_k!}\left(\frac{1}{k}\right)^{n-2} = \\ = k^2 \sum \{n_1^2 - n_1 - n_1 n_2\} f_n, \quad (30.56)$$

где $f_n = \frac{n!}{n_1!\, n_2! \ldots n_k!} k^{-n}$ и суммирование всюду, где это не

указано, производится по критической области $X^2 \geqslant c_\alpha$, которая в силу выражения (30.44) для X^2 имеет вид

$$\sum_{i=1}^{k} n_i^2 \geqslant b_\alpha.$$

Величина $\frac{1}{\alpha}\sum n_1^2 f_n$ есть усреднение n_1^2 по критической области, и оно должно превосходить среднее значение n_1^2 по всей совокупности, равное $\frac{n}{k}\left(1-\frac{1}{k}+\frac{n}{k}\right)$ в силу (30.45). Таким образом,

$$\sum n_1^2 f_n = \alpha\frac{n}{k}\left(1-\frac{1}{k}+\frac{n}{k}\right)+d, \tag{30.57}$$

где $d>0$. Точно такими же рассуждениями, поскольку n_1 положительно, мы получаем $\frac{1}{\alpha}\sum n_1 f_n > \frac{n}{k}$, или

$$\sum n_1 f_n = \alpha(n/k)+e, \tag{30.58}$$

где $e>0$. Кроме того, мы, очевидно, имеем в (30.57) и (30.58)

$$d>e. \tag{30.59}$$

В силу симметрии мы имеем

$$\sum n_1 n_2 f_n = \sum\left\{\frac{1}{k(k-1)}\left(n^2-\sum_i n_i^2\right)\right\} f_n = \frac{\alpha n^2}{k(k-1)} - \frac{1}{k-1}\sum n_1^2 f_n. \tag{30.60}$$

Пользуясь (30.57)—(30.60), получаем из (30.56)

$$\frac{1}{k^2}\left[\frac{\partial^2 P}{\partial\theta_1^2}-\frac{\partial^2 P}{\partial\theta_1\,\partial\theta_2}\right] > \frac{k}{k-1}\cdot\alpha\frac{n}{k}\left(1-\frac{1}{k}+\frac{n}{k}\right)-\alpha\frac{n}{k}-\alpha\frac{n^2}{k(k-1)}=0. \tag{30.61}$$

Таким образом, второй член в правой части (30.54) положителен. Члены высших порядков, отброшенные в (30.54), содержат θ_i в третьей и более высоких степенях и вблизи H_0 будут, следовательно, меньше по модулю, чем член второго порядка. Таким образом, $P \geqslant \alpha$ вблизи H_0 и критерий с равными вероятностями локально несмещен, что служит рекомендацией для этой процедуры образования классов, так как никаких результатов такого рода для критерия X^2 в общем случае неизвестно.

Предельная функция мощности

30.27 Предположим, что, как и при обсуждении асимптотической относительной эффективности в главе 25, мы позволяем H_1 приближаться к H_0 при возрастании n со скоростью, достаточной, чтобы мощность не приближалась к 1. Положим

$p_{1i} - p_{0i} = c_i n^{-1/2}$, где c_i фиксированы. Тогда X^2 имеет асимптотически нецентральное χ^2-распределение с $k - s - 1$ степенями свободы (когда s параметров оцениваются посредством мультиномиальных МП-оценок) и параметром нецентральности

$$\lambda = \sum_{i=1}^{k} \frac{c_i^2}{p_{0i}} = n \sum_{i=1}^{k} \frac{(p_{1i} - p_{0i})^2}{p_{0i}}. \tag{30.62}$$

Этот результат, впервые объявленный Эйзенхартом (1938), сразу следует из представления, указанного в **30.9—10**; его доказательство предоставляется читателю в качестве упражнения 30.4. Теперь, применяя лемму Неймана — Пирсона (22.6) к (24.18), получаем, что наилучшая критическая область для проверки $\lambda = 0$ представляет собой верхний «хвост» распределения X^2.

Аппроксимация нецентрального χ^2 в **24.5** позволяет нам приближенно оценить мощность X^2-критерия. Она дается тогда интегралом (24.30). Для $\alpha = 0{,}05$ могут быть использованы таблицы Патнайка, описанные в **24.5**.

Пример 30.3

Мы можем проиллюстрировать использование предельной функции мощности, возвращаясь к задаче примера 30.2 и исследуя влияние удвоения k на мощность критерия с равными вероятностями. Чтобы облегчить использование таблиц из *Biometrika Tables*, мы в действительности выбираем четыре класса с несколько неравными вероятностями:

Значения	p_{0i}	p_{1i}	$(p_{1i}-p_{0i})^2$	$(p_{1i}-p_{0i})^2/p_{0i}$
0—0,3	0,259	0,104	0,0240	0,0927
0,3—0,7	0,244	0,190	0,0029	0,0119
0,7—1,4	0,250	0,282	0,0010	0,0040
1,4 и больше	0,247	0,424	0,0313	0,1267
				$0{,}2353 = \frac{\lambda}{n}$

В этой таблице p_{0i} получены, как и раньше, из гамма-распределения с параметром 1, а p_{1i} — из гамма-распределения с параметром 1,5. Для этих четырех классов и $n = 50$, как в примере 30.2, мы находим, что параметр нецентральности (30.62) равен $\lambda = 0{,}2353 \cdot 50 = 11{,}8$. Для X^2 с 3 степенями

свободы, когда $\alpha = 0{,}05$, это дает мощность 0,83 по таблице Патнайка.

Предположим теперь, что мы образуем восемь классов, разбивая каждый из имеющихся классов на два, с новыми p_{0i}, равными настолько, насколько это удобно для использования таблиц. Мы находим:

Значения	p_{0i}	p_{1i}	$(p_{1i}-p_{0i})^2$	$(p_{1i}-p_{0i})^2/p_{0i}$
0—0,15	0,139	0,040	0,0098	0,0705
0,15—0,3	0,120	0,064	0,0031	0,0258
0,3 —0,45	0,103	0,071	0,0010	0,0097
0,45—0,7	0,141	0,119	0,0005	0,0035
0,7 —1,0	0,129	0,134	0,0000	0,0002
1,0 —1,4	0,121	0,148	0,0007	0,0058
1,4 —2,1	0,125	0,183	0,0034	0,0272
2,1 и больше	0,122	0,241	0,0142	0,1163
				$0{,}2590 = \frac{\lambda}{n}$

Для $n = 50$ теперь имеем $\lambda = 13{,}0$ с семью степенями свободы, и мощность при $\alpha = 0{,}05$ приближенно равна 0,75 по таблице Патнайка. Удвоение k увеличило λ, но лишь немного. Мощность же на самом деле уменьшилась, так как при данном λ центральное и нецентральное χ^2-распределения сближаются при возрастании числа степеней свободы (см. упражнение 24.3), и здесь это влияет сильнее, чем увеличение λ. Однако n здесь слишком мало для того, чтобы делать надежные заключения о мощности по предельной функции мощности, и, возможно, следует сделать вывод, что удвоение k повлияло на мощность очень мало.

Выбор k в критерии с равными вероятностями *)

30.28 Манном и Вальдом (1942) был получен следующий результат относительно выбора k. Пусть для проверки простой гипотезы (30.1) с непрерывной ф. р. $F_0(x)$ по выборке объема n применяется критерий X^2 с k равновероятными классами. Обозначим $f(n, k, \Delta)$ минимум мощности критерия против класса альтернативных ф. р.

$$H_\Delta = \{F(x): \sup_{-\infty < x < \infty} |F(x) - F_0(x)| \geqslant \Delta\}.$$

*) В оригинале п. **30.28** содержит вывод формулы (30.72) (см. примечание на стр. 585), которому трудно придать какой-либо смысл. Приводимый здесь текст п. **30.28** написан Д. М. Чибисовым. (*Прим. ред.*)

Для данных n и Δ функция $f(n, k, \Delta)$ достигает максимального значения, скажем $f(n, \Delta)$, при некотором $k = k(n, \Delta)$. Выберем $\Delta = \Delta_n$ так, чтобы значение $f(n, \Delta_n)$ равнялось заданному P_0; для определенности принимается $P_0 = 1/2$. Оптимальным считается число классов $k_n = k(n, \Delta_n)$, т. е. число k, доставляющее максимум функции $f(n, k, \Delta)$ при таком Δ, для которого этот максимум равен 1/2. Пусть $G(x)$ — стандартная нормальная ф. р. и d_α — решение уравнения $G(d_\alpha) = 1 - \alpha$. Теорема Манна и Вальда утверждает, что при $n \to \infty$

$$k_n \sim 4\sqrt[5]{2}\,(n/d_\alpha)^{2/5}. \tag{30.63}$$

Доказательство проводится по следующей схеме. Можно считать, что $F_0(x) = x$ при $0 \leqslant x \leqslant 1$ (равномерное распределение на $(0, 1)$), а все альтернативные распределения также сосредоточены на $(0, 1)$ (см. **30.36**). Если наблюдениям соответствует ф. р. $F(x)$, то X^2 имеет приближенно нецентральное χ^2-распределение с $\nu = k - 1$ степенями свободы и параметром нецентральности

$$\lambda = nk\sum_{i=1}^{k}\left(p_{1i} - \frac{1}{k}\right)^2 = n\left(k\sum_i p_{1i}^2 - 1\right) \tag{30.64}$$

(см. (30.62)), где $p_{1i} = F(i/k) - F((i-1)/k)$. При этом (см. (24.19))

$$\mathbf{M}X^2 \approx \nu + \lambda, \quad \mathbf{D}X^2 \approx 2(\nu + 2\lambda). \tag{30.65}$$

Оптимальное k_n должно возрастать с ростом n, и тогда статистика X^2 асимптотически нормальна с параметрами (30.65). Поэтому мощность критерия равна приближенно

$$\mathbf{P}\{X^2 > \nu + d_\alpha\sqrt{2\nu} \mid H_1\} \approx G\left(\frac{\lambda - d_\alpha\sqrt{2\nu}}{\sqrt{2(\nu + 2\lambda)}}\right) = P(k, \lambda). \tag{30.66}$$

Нетрудно проверить, что $(\partial/\partial\lambda)P(k, \lambda) > 0$, и, таким образом, для нахождения $f(n, k, \Delta)$ нужно найти ф. р. из H_Δ, которой соответствует наименьшее λ. Рассмотрим эквивалентную задачу: при фиксированном λ найти ф. р. $F(x)$, для которой $\sup_{0 \leqslant x \leqslant 1} |F(x) - x| = \max$. Поскольку λ определяется только значениями $F(i/k)$, $i = 0, 1, \ldots, k$, рассмотрим сначала

$$\max_i \left|F\left(\frac{i}{k}\right) - \frac{i}{k}\right| = \max_i \left|\sum_{l=1}^{i}\left(p_{1l} - \frac{1}{k}\right)\right|. \tag{30.67}$$

Пусть для простоты k четно. Тогда (30.67) достигает максимума при условиях $\sum(p_{1i} - (1/k))^2 = \lambda/nk$ и $\sum(p_{1i} - (1/k)) = 0$,

когда $p_{1i} - (1/k)$ равны, скажем, δ при $i = 1, \ldots, k/2$ и равны $-\delta$ при $i = k/2 + 1, \ldots, k$. Согласно (30.64) $\delta = \pm k^{-1} (\lambda/n)^{1/2}$; в дальнейшем для определенности будем считать, что $\delta > 0$. Это означает, что разность $F(i/k) - (i/k)$ с ростом i возрастает линейно при $i = 0, \ldots, k/2$ до значения $k\delta/2 = (\lambda/n)^{1/2}/2$ и затем линейно убывает до нуля при $i = k/2 + 1, \ldots, k$. Теперь, не меняя значений $F(i/k)$, можно добиться наибольшего отклонения $F(x)$ от x, если положить $F(x) = F(i/k)$ при $(i-1)/k < x < i/k$. Тогда $\sup |F(x) - x|$ достигается при $x = \left(\frac{1}{2} - \frac{1}{k}\right) + 0$ и равен $\Delta = \frac{1}{2}\sqrt{\frac{\lambda}{n}} + \frac{1}{k}$. Отсюда находим, что $\lambda = 4n(\Delta - (1/k))^2$ и

$$f(n, k, \Delta) \approx G\left(\frac{4n(\Delta - (1/k))^2 - d_\alpha \sqrt{2(k-1)}}{\sqrt{2[k - 1 + 8n(\Delta - (1/k))^2]}}\right). \qquad (30.68)$$

Для нахождения k_n нужно максимизировать по k аргумент функции G в (30.68); это асимптотически эквивалентно максимизации его числителя. Дифференцируя по k, получаем

$$8n\left(\Delta - \frac{1}{k}\right)\frac{1}{k^2} - \frac{d_\alpha}{\sqrt{2(k-1)}} = 0. \qquad (30.69)$$

Кроме того, условие $P(n, \Delta_n) = P_0 = 1/2$ дает

$$4n\left(\Delta - \frac{1}{k}\right)^2 - d_\alpha \sqrt{2(k-1)} = 0. \qquad (30.70)$$

Из (30.69) и (30.70), пренебрегая различием между k и $k-1$, получаем (30.63). При произвольном P_0 можно аналогичным образом получить из (30.68) и (30.69), что *)

$$k_n \sim 4\sqrt[5]{2}\left(n\frac{d_\alpha + G^{-1}(P_0)}{d_\alpha^2}\right)^{2/5}. \qquad (30.71)$$

Следует заметить, что результат Манна и Вальда получен из условия максимизации мощности против самых неблагоприятных альтернатив при заданном расхождении функций распределения. Если же иметь ввиду только достаточно «гладкие»

*) В оригинале для оптимального значения k получена формула

$$k \sim b\left\{\frac{2^{1/2}(n-1)}{d_\alpha + G^{-1}(P_0)}\right\}^{2/5}, \qquad (30.72)$$

где коэффициент b связан с характером поведения альтернатив. (*Прим. ред.*)

альтернативы, то оптимальное число k будет значительно меньше. Качественно это можно объяснить следующим образом. Пусть распределения при H_0 и H_1 задаются плотностями $p_0(x)$ и $p_1(x)$ соответственно, причем $p_1(x)/p_0(x) = 1 + n^{-1/2}c(x)$. Тогда (как и в **30.27**) $p_{1i} - p_{0i} = c_i n^{-1/2}$, где $c_i = \int\limits_{\Delta_i} c(x)\, p_0(x)\, dx$ и Δ_i есть i-й класс-интервал. Параметр нецентральности (30.62) будем обозначать λ_k. По неравенству Коши — Буняковского имеем

$$c_i^2 \leqslant p_{0i} \int\limits_{\Delta_i} c^2(x)\, p_0(x)\, dx,$$

и, следовательно, $\lambda_k \leqslant \lambda_\infty \equiv \int c^2(x)\, p_0(x)\, dx$. Предполагая последний интеграл конечным, нетрудно показать, что $\lambda_k \to \lambda_\infty$, если $k \to \infty$ и все $p_{0i} \to 0$. Пусть $P(k, \lambda_k)$ — мощность критерия X^2 с k классами, когда параметр нецентральности равен λ_k. Когда k невелико, нормальная аппроксимация, дающая (30.67), недействительна, но характер поведения $P(k, \lambda)$ остается тем же: $P(k, \lambda)$ возрастает по λ и убывает по k. Поэтому $P(k, \lambda_k)$, будучи ограничена сверху функцией $P(k, \lambda_\infty)$, достигает максимума при некотором k, зависящем от скорости приближения λ_k к λ_∞ (и не зависящем от n). Эта скорость определяется гладкостью функции $c(x)$, и при достаточно гладких альтернативах максимум $P(k, \lambda_k)$ может достигаться при небольших k. Ниже приводятся данные из примеров 30.3, 30.4, дополненные результатами вычислений для $k = 2$, 3 и 5, которые показывают, что оптимальное значение k в этом случае равно 3 (или $k \leqslant 4$ ввиду возможных погрешностей вычислений и χ^2-аппроксимации при $n = 50$).

k	2	3	4	5	8	15
λ_k	8,7	10,9	11,8	12,4	13,0	13,7
P	0,83	0,84	0,83	0,81	0,75	0,64

Эти данные показывают также быстрое приближение λ_k к λ_∞, поскольку здесь $\lambda_\infty = 50[(4/\pi) - 1] = 13{,}66$ (нарушение неравенства $\lambda_{15} \leqslant \lambda_\infty$ в данном случае объясняется ошибками округления при вычислениях в примере 30.4). По-видимому, этот пример представляет крайний случай малости оптимального k, но он показывает, что для представляющих обычно ин-

терес альтернатив теорема Манна и Вальда дает сильно завышенное значение k.

30.29 Мы заключаем, что k в случае равных вероятностей следует увеличивать пропорционально $n^{2/5}$ и что k должно быть меньше, когда нас интересует область большой мощности (когда $G^{-1}(P_0)$ велико), чем когда нас интересует область малых значений мощности *).

Формула (30.63) приводит к значениям k, гораздо бóльшим, чем применяемые обычно. При увеличении n в $4\sqrt{2}$ раз k удваивается. Когда $n = 200$, $k = 31$ при $\alpha = 0{,}05$ и $k = 27$ при $\alpha = 0{,}01$, — эти значения близки к наименьшим значениям k, при которых вообще допустима нормальная аппроксимация, применяемая в наших рассуждениях (а также у Манна и Вальда). В этом случае Манн и Вальд рекомендуют использовать (30.63), когда $n \geqslant 450$ при $\alpha = 0{,}05$ и когда $n \geqslant 300$ при $\alpha = 0{,}01$. Можно видеть, что гипотетическая ожидаемая частота каждого класса, равная n/k, возрастает, как $n^{3/5}$, и равна примерно 6 и 8 соответственно, когда $n = 200$, $\alpha = 0{,}05$ и $\alpha = 0{,}01$.

Уилльямс (1950) сообщает, что оптимальное значение k Манна и Вальда может быть уменьшено вдвое без серьезной потери мощности в точке 0,50. Но следует помнить, что n и k должны быть достаточно большими, чтобы (30.63) давало хорошие результаты. Пример 30.4 иллюстрирует это положение, которое подтверждается также вычислениями, проделанными Хэмдэном (1963) для критериев среднего значения в нормальном распределении.

Пример 30.4

Рассмотрим снова задачу примера 30.3. Мы получили там для мощности значение, близкое к 0,8. Из таблицы нормального распределения $G^{-1}(0{,}8) = 0{,}84$. При $b = 4$, $\alpha = 0{,}05$, $\lambda_\alpha = 1{,}64$ формула (30.72) дает для оптимального значения k вблизи этой точки значение **)

$$k = 4\left\{\frac{2^{1/2}(n-1)}{2{,}48}\right\}^{2/5} = 3{,}2\,(n-1)^{2/5}.$$

При $n = 50$ это дает приблизительно $k = 15$.

Предположим теперь, что мы пользуемся таблицами *Biometrika Tables* для построения группировки по 15 классам с ве-

*) Этот вывод основан на (30.72). На самом деле имеет место противоположный характер зависимости k от P_0, см. (30.71). (*Прим. ред.*)

**) См. примечание на стр. 585. Применение (30.71) с $P_0 = 0{,}5$ и $P_0 = 0{,}8$ дало бы соответственно $k = 18$ и $k = 21$. (*Прим. ред.*)

роятностями p_{0i}, приближенно равными между собой настолько, насколько это удобно. Мы находим:

Значения	p_{0i}	p_{1i}	$(p_{1i}-p_{0i})^2/p_{0i}$	Значения	p_{0i}	p_{1i}	$(p_{1i}-p_{0i})^2/p_{0i}$
0—0,05	0,049	0,008	0,034	0,75—0,90	0,065	0,067	0,000
0,05—0,15	0,090	0,032	0,037	0,90—1,1	0,074	0,083	0,000
0,15—0,20	0,042	0,020	0,012	1,1—1,3	0,060	0,075	0,004
0,20—0,30	0,078	0,044	0,015	1,3—1,6	0,071	0,095	0,008
0,30—0,40	0,071	0,047	0,008	1,6—2,0	0,067	0,101	0,018
0,40—0,50	0,063	0,048	0,004	2,0—2,7	0,068	0,116	0,034
0,50—0,65	0,085	0,072	0,002	2,7 и больше	0,067	0,145	0,098
0,65—0,75	0,050	0,047	0,000				
							$0{,}274=\lambda/n$

Здесь $\lambda = 13{,}7$ и таблица Патнайка дает мощность 0,64 при 14 степенях свободы. Снова λ увеличилось, но мощность уменьшилась из-за увеличения k. Мы не получили здесь оптимума. При большом k (и, следовательно, большом n) увеличение k не подавило бы таким образом эффект увеличения λ.

30.30 Мы не должны брать k слишком большим, поскольку мультиномиальное распределение не может хорошо аппроксимироваться многомерным нормальным, когда np_{0i} очень малы. Грубое правило, которым обычно пользуются, состоит в том, что ожидаемые частоты (np_{0i}) должны быть не меньше 5. Для этого правила, по-видимому, нет общих теоретических оснований, и относительно него стоит сделать следующие два замечания. Если распределение при H_0 унимодально и обычным образом используются классы одинакового размера, то малые ожидаемые частоты возникают только на «хвостах». Кокрэн (1952, 1954) рекомендует принять гибкий подход; он проверил, что для одной или двух ожидаемых частот допустимо уменьшение до 1 и даже ниже, если X^2 имеет по крайней мере 6 степеней свободы, без ущерба для критерия с $\alpha = 0{,}05$ или 0,01.

В случае равных вероятностей все ожидаемые частоты будут равны. Слэктер (1966) показал, что даже для равных ожидаемых частот, меньших единицы, аппроксимация остается хорошей. Процедура Манна — Вальда из **30.29** ведет к ожидаемым частотам, всегда бóльшим, чем 5, при $n \geqslant 200$. Интересно отметить, что в примерах 30.3, 30.4 применение этого ограничения исключило бы процедуру с 15 классами и была бы приемлема более мощная процедура с 8 классами, при которой ожидаемые частоты имеют значения от 5 до 7.

Гёфдинг (1965) показал, что если при проверке простой гипотезы H_0 (случай сложной гипотезы менее ясен) k остается фиксированным, тогда

как $\alpha \to 0$ подходящим образом при $n \to \infty$, то критерий отношения правдоподобия будет более мощным, чем X^2-критерий. Этот результат не выполняется, если k возрастает с n, например, как в (30.72) в случае равных вероятностей.

Отметим, что асимптотический характер теории распределения X^2 не является недостатком в практических применениях, поскольку мы обычно не занимаемся проверкой согласия, не имея большой выборки.

Рекомендации для критерия X^2

30.31 Мы резюмируем предыдущее обсуждение несколькими практическими рекомендациями:

(1) Если проверяемое распределение табулировано, применять классы с равными или примерно равными вероятностями.

(2) Определять число классов, когда n превышает 200, приближенно по формуле (30.72) с b между 2 и 4.

(3) Если требуется оценивать параметры, пользоваться обычными МП-оценками в интересах эффективности*), но помнить, что при этом происходит частичное возникновение степеней свободы (**30.19**), так что критические значения должны быть повышены; если используются мультиномиальные МП-оценки, то такого повышения не требуется.

Развитая выше теория неприменима вовсе, если по данным, используемым для проверки согласия, оценивается функциональная форма распределения (а не только параметры).

30.32 Помимо трудностей, связанных с критериями X^2, которые мы уже обсуждали и которые не очень серьезны, эти критерии подвергались критике в двух направлениях. В обоих случаях критика касается мощности критерия. Во-первых, тот факт, что существенным шагом является переход к задаче о мультиномиальном распределении, влечет необходимость группировки наблюдений по классам. Такая группировка должна вести к потере информации, понимаемой в общем широком смысле, и мы подозреваем, что эта потеря будет наибольшей, когда проверяется согласие с непрерывным распределением. Во-вторых, тот факт, что статистика X^2 основана на квадратах отклонений наблюденных частот от гипотетических, влечет отсутствие чувствительности к характеру поведения знаков этих отклонений, который, очевидно, содержит информацию. Первая из этих критик более радикальна, поскольку она должна, очевидно, привести к поиску других статистик критериев, которые заменили бы X^2, и мы рассмотрим такие критерии после обсуждения второй критики.

*) См. примечание на стр. 570. (*Прим. ред.*)

Знаки отклонений

30.33 Рассмотрим, какого поведения отклонений наблюденных частот от гипотетических следует ожидать в некоторых простых случаях. Предположим, что простая гипотеза задает непрерывное унимодальное распределение с параметрами сдвига и масштаба, равными, скажем, среднему и стандартному отклонению, и предположим, что гипотетическое среднее слишком велико. Для любого набора k классов вероятности p_{0i} будут очень малы при малых значениях варианты, а затем очень велики, как показано на рис. 30.1. Поскольку при больших выборках наблюдаемые относительные частоты будут сходиться по вероятности к истинным вероятностям, характерным поведением наблюдаемых отклонений будет серия положительных отклонений, за которой следует серия отрицательных. Если гипотетическое среднее слишком мало, то характер поведения будет обратным.



Рис. 30.1. Гипотетическое и истинное распределения, отличающиеся сдвигом.

Предположим теперь, что гипотетическое значение параметра масштаба слишком мало. Тогда картина будет такой, как на рис. 30.2. Мы видим, что характерным поведением отклонений при больших выборках теперь будет серия положительных отклонений, за ней — серия отрицательных и за ней — серия положительных. Если гипотетический параметр масштаба слишком велик, то все эти знаки меняются на обратные.

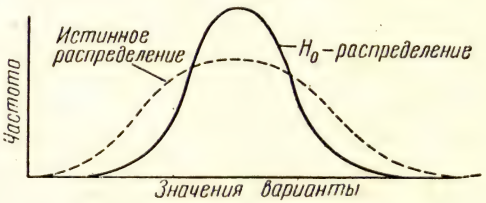


Рис. 30.2. Гипотетическое и истинное распределения, отличающиеся масштабом.

Разумеется, мы не применяем критерий X^2, зная, что возможны лишь изменения в сдвиге и масштабе, так как тогда мы можем найти более мощные критерии. Однако, когда возможна ошибка как в параметре сдвига, так и в параметре масштаба, рис. 30.3 показывает, что ситуация остается по существу той же самой; мы по-прежнему будем иметь три (или в более сложных случаях несколько больше) серии знаков отклонений. Более общим образом, если истинные значения параметров отличаются от гипотетических или если форма распределения отличается

«гладко» от гипотетической формы, мы ожидаем, что знаки отклонений будут собираться в подобного рода группы, а не распределяться случайно, как это было бы, если бы гипотетические вероятности совпадали с истинными.

30.34 Это наблюдение наводит на мысль дополнить критерий X^2 критерием числа серий знаков наблюдений, в котором малые числа образуют критическую область. Элементарная теория серий, необходимая для этой цели, дается как упражнение 30.8. Однако для ее корректного применения мы должны сначала исследовать взаимосвязь между критерием «серий» и критерием X^2. Ф. Дэвид (1947), Сил (1948) и Фрэзер (1950) показали, что, когда справедлива H_0, эти критерии асимптотически независимы (см. упражнение 30.7) и что при проверке простой гипотезы все последовательности знаков равновероятны, так что теория из упражнения 30.8 может быть скомбинирована с критерием X^2, как указано в упражнении 30.9.



Рис. 30.3. Гипотетическое и истинное распределения, отличающиеся сдвигом и масштабом.

Дополнение критерием «серий», по-видимому, полезно для повышения чувствительности, когда проверяется простая гипотеза, как было показано выше, при иллюстративном обсуждении. Для сложной гипотезы, представляющей в связи с критериями согласия особый интерес, когда все параметры должны оцениваться по выборке, оно не имеет практического значения, так как последовательности знаков отклонений, хотя и независимы от X^2, уже не равновероятны, как было в случае простой гипотезы, и теория из упражнения 30.8, следовательно, неприменима (см. Фрэзер (1950)).

Другие критерии согласия

30.35 Мы обратимся теперь к обсуждению альтернативных критериев согласия. Поскольку они направлены на то, чтобы избежать потери информации от группировки, присущей критерию X^2, для них не проходит упрощение, связанное с переходом к мультиномиальному распределению, и мы должны ожидать, что их теория будет более трудной. Прежде чем обсуждать наиболее важные критерии в отдельности, мы отметим их общие черты.

Заметим, что, когда статистика X^2 применяется для проверки простой гипотезы, она распределена асимптотически как χ^2_{k-1}, *какова бы ни была эта простая гипотеза*, хотя ее точное распределение и зависит от распределения, задаваемого гипотезой. Ясно, что такой результат достигается за счет привлечения мультиномиального распределения и его сходимости к нормальному. Кроме того, то же самое справедливо и в случае сложной гипотезы, когда применяются мультиномиальные МП-оценки, в этом случае $X^2 \to \chi^2_{k-s-1}$, *какова бы ни была сложная гипотеза*, хотя здесь еще более ясно, что точное распределение X^2 зависит от рассматриваемой сложной гипотезы. Когда используются другие оценки (даже когда это эффективные обычные МП-оценки), эти приятные асимптотические свойства не выполняются: даже асимптотическое распределение X^2 теперь зависит от характеристических корней матрицы (30.37), которые в общем случае являются функциями как гипотетического распределения, так и значений параметров θ.

Выразим эти результаты в следующей форме: в первых двух случаях мы говорим, что распределение X^2 асимптотически *свободно от распределения* (т. е. на него не влияет форма гипотетического распределения или значения параметров), в то время как в третьем случае оно не будет ни асимптотически свободным от распределения, ни даже *свободным от параметров* (т. е. свободным ст влияния параметров F_0, хотя и зависящим от вида распределения F_0).

30.36 Мы увидим, что во всех наиболее важных критериях согласия прямо или косвенно используется *вероятностное интегральное преобразование*, которое нам встречалось в различных ситуациях (например, **1.27**, **24.11**) как средство для преобразования любого известного непрерывного распределения в прямоугольное распределение на интервале (0, 1). В наших тепе решних обозначениях, если мы имеем простую гипотезу, задающую ф. р. $F_0(x)$, которой соответствует функция плотности $f_0(x)$, то величина $y = \int_{-\infty}^{x} f_0(u)\,du = F_0(x)$ имеет равномерное распределение на (0, 1). Поэтому, если мы имеем n наблюдений x_i и преобразуем их в y_i посредством вероятностного интегрального преобразования при известной $F_0(x)$, а затем используем функцию от y_i для проверки отклонения y_i от равномерности, то распределение статистики критерия будет свободным от распределения не только асимптотически, но и при каждом n.

Когда распределение задается сложной гипотезой, скажем $F_0(x|\theta_1, \theta_2, \ldots, \theta_s)$, где s параметров θ подлежат оценке, мы должны выбрать для этого s функций $t_1, \ldots, t_s$ от x_i.

Преобразованные случайные величины теперь равны

$$y_i = \int_{-\infty}^{x_i} f_0(u \mid t_1, t_2, \ldots, t_s)\, du,$$

но они не независимы и не имеют равномерного распределения, а их распределение, вообще говоря, будет зависеть как от гипотетического распределения F_0, так и от истинных значений его параметров, как было показано Дэвидом и Джонсоном (1948). Однако (см. упражнение 30.10) если F имеет только параметры сдвига и масштаба и их оценки в должном смысле инвариантны, то распределение y_i будет зависеть от формы F, но не от его параметров. Следовательно, при конечных n никакая статистика критерия, основанная на y_i, не может быть свободной от распределения в случае сложной гипотезы согласия (хотя она может быть свободной от параметров, если имеются только параметры сдвига и масштаба). Разумеется, такая статистика по-прежнему может быть асимптотически свободной от распределения.

«Гладкие» критерии Неймана — Бартона

30.37 Первыми из критериев согласия, которые мы рассмотрим в качестве альтернатив к X^2, будут так называемые «гладкие» критерии, разработанные впервые Нейманом (1937a), рассматривавшим, насколько нам известно, только простую гипотезу. При заданной H_0: $F(x) = F_0(x)$ мы преобразуем n наблюдений x_i, как в **30.36**, с помощью вероятностного интегрального преобразования

$$y_i = \int_{-\infty}^{x_i} f_0(u)\, du = F_0(x_i), \qquad i = 1, 2, \ldots, n, \tag{30.73}$$

и получаем n независимых наблюдений, распределенных равномерно на интервале (0, 1), когда справедлива H_0. В качестве альтернатив к H_0 мы задаем отклонения от равномерности распределения y_i, которые тем не менее остаются независимыми. Нейман построил систему распределений, которые могли бы гладко изменяться, отклоняясь от H_0-распределения (равномерного), и зависели бы от нескольких параметров. (Именно эта «гладкость» альтернатив была перенесена на название самих критериев.) Альтернативы, которые рассматривал Нейман, задавались для всех y_i функцией плотности

$$f(y \mid H_k) = c(\theta_1, \theta_2, \ldots, \theta_k) \exp\left\{1 + \sum_{r=1}^{k} \theta_r \pi_r(y)\right\}, \tag{30.74}$$

$$0 \leqslant y \leqslant 1, \qquad k = 1, 2, 3, \ldots,$$

где c — константа, которая обеспечивает равенство единице интеграла от (30.74), а $\pi_r(y)$ — полиномы Лежандра, линейно преобразованные так, чтобы они были ортогональны на интервале (0, 1) *). Если положить $z = y - \frac{1}{2}$, то этими полиномами до четвертого порядка будут

$$\left.\begin{aligned} \pi_0(z) &\equiv 1, \\ \pi_1(z) &= 3^{1/2} \cdot 2z, \\ \pi_2(z) &= 5^{1/2} \cdot \left(6z^2 - \tfrac{1}{2}\right), \\ \pi_3(z) &= 7^{1/2} \cdot (20z^3 - 3z), \\ \pi_4(z) &= 3 \cdot \left(70z^4 + 15z^2 + \tfrac{3}{8}\right). \end{aligned}\right\} \tag{30.75}$$

30.38 Теперь задача состоит в том, чтобы найти статистику критерия для проверки H_0 против H_k. Можно видеть, что если переписать семейство (30.74) в виде

$$f(y \mid H_k) = c(\theta) \exp\left\{\sum_{r=0}^{k} \theta_r \pi_r(y)\right\}, \qquad 0 \leqslant y \leqslant 1, \qquad k = 0, 1, 2, \ldots, \tag{30.76}$$

считая $\theta_0 \equiv 1$, то оно включает в себя H_0. Мы хотим проверить простую гипотезу

$$H_0\colon\ \theta_1 = \theta_2 = \ldots = \theta_k = 0, \tag{30.77}$$

или, эквивалентно,

$$H_0\colon\ \sum_{r=1}^{k} \theta_r^2 = 0, \tag{30.78}$$

---

*) Полиномы Лежандра, скажем $L_r(z)$, обычно определяются формулой

$$L_r(z) = (r!\, 2^r)^{-1} \frac{d^r}{dz^r}\{(z^2 - 1)^r\}$$

и удовлетворяют условиям ортогональности

$$\int_{-1}^{1} L_r(z)\, L_s(z)\, dz = \begin{cases} 0, & r \neq s, \\ \dfrac{2}{2r+1}, & r = s. \end{cases}$$

Следовательно, чтобы сделать их ортогональными на $\left(-\frac{1}{2}, \frac{1}{2}\right)$, мы определяем полиномы $\pi_r(z)$ как $\pi_r(z) = (2r+1)^{1/2} L_r(2z)$. Мы могли бы теперь перейти к интервалу (0, 1), положив $y = z + \frac{1}{2}$. Но удобнее, как в тексте, иметь дело с $z = y - \frac{1}{2}$.

против сложной альтернативы, состоящей в отрицании этого. Мы видим, что (30.76) образует экспоненциальное семейство, линейное по θ_r и π_r. Функция правдоподобия n независимых наблюдений равна

$$L(y\,|\,\theta) = \{c(\theta)\}^n \exp\left\{\sum_{r=0}^{k} \theta_r \sum_{i=1}^{n} \pi_r(y_i)\right\}. \qquad (30.79)$$

Выражение (30.79), очевидно, распадается на k сомножителей, так что каждая из статистик $t_r = \sum_{i=1}^{n} \pi_r(y_i)$ достаточна для θ_r, и мы, следовательно, можем ограничиться функциями от t_r при отыскании статистики критерия. Когда мы имели дело с линейными функциями от θ_r в **23.27—32**, мы видели, что та же функция от t_r дает РНМН критерий. Здесь нас интересует сумма квадратов параметров, и кажется разумным использовать соответствующую функцию от t_r, т. е. $\sum_{r=1}^{k} t_r^2$, в качестве статистики критерия, хотя мы и не можем ожидать, что она будет иметь такое же сильное свойство оптимальности. Это и была, с точностью до константы, статистика, предложенная Нейманом (1937а), который для обоснования ее выбора использовал асимптотические рассуждения. Пирсон (1938) показал, что при больших выборках эта статистика эквивалентна критерию отношения правдоподобия для гипотезы (30.78). Мы положим $u_r = n^{-1/2} t_r$; статистика критерия тогда равна *)

$$p_k^2 = \sum_{r=1}^{k} u_r^2 = \frac{1}{n} \sum_{r=1}^{k} \left\{\sum_{i=1}^{n} \pi_r(y_i)\right\}^2. \qquad (30.80)$$

30.39 Так как $u_r = n^{-1/2} \sum_{i=1}^{n} \pi_r(y_i)$, то по центральной предельной теореме u_r распределены асимптотически нормально со средними и дисперсиями, равными в силу (30.79)

$$\mathbf{M}(u_r) = n^{1/2}\mathbf{M}\{\pi_r(y)\} = n^{1/2}\theta_r, \qquad (30.81)$$

$$\mathbf{D}(u_r) = \mathbf{D}\{\pi_r(y)\} = 1, \qquad (30.82)$$

и u_r некоррелированы, так как π_r ортогональны. Таким образом, статистика (30.80) асимптотически есть сумма квадратов независимых нормальных случайных величин с единичными дисперсиями, у которых все средние значения равны нулю тогда и

*) Эту статистику обычно обозначают ψ_k^2; мы отказываемся от этого обозначения в соответствии с нашим соглашением употреблять латинские буквы для статистик и греческие для параметров.

только тогда, когда справедлива H_0. Следовательно, p_k^2 асимптотически имеет нецентральное χ^2-распределение с k степенями свободы и параметром нецентральности, равным согласно (30.81)

$$\lambda = n \sum_{r=1}^{k} \theta_r^2. \tag{30.83}$$

Отсюда сразу следует, что p_k^2 определяет состоятельный (и, по **24.17**, асимптотически несмещенный) критерий, как показал Нейман (1937а). Дэвид (1939) нашел, что при гипотезе H_0 простейшие статистики p_1^2 и p_2^2 достаточно хорошо аппроксимируются (центральными) χ^2-распределениями с 1 и 2 степенями свободы соответственно при $n \geqslant 20$.

Формулировка альтернативных гипотез

30.40 Выбор k, порядка системы альтернатив (и числа параметров, которыми выражается отклонение от H_0), должен быть сделан до построения критерия. Ясно, что мы не хотим иметь больше параметров, чем требуется для интересующих нас альтернатив, так как они будут «разжижать» критерий. К сожалению, при проверке согласия часто не имеют в виду достаточно определенных альтернатив. Это очень реальная трудность, которую можно сравнить с выбором числа классов в X^2-критерии. В последнем случае мы нашли, что этот выбор может производиться в зависимости лишь от объема выборки и размера критерия; для нашей нынешней же ситуации еще не имеется достаточно ясного руководства.

30.41 В первой из серии статей Бартона (1953, 1955, 1956), на работах которого базируются следующие разделы, рассматривалась несколько иная общая система альтернатив. Вместо (30.76) он определяет

$$f(y \mid H_k) = \sum_{r=0}^{k} \theta_r \pi_r(y), \qquad 0 \leqslant y \leqslant 1, \quad k = 0, 1, 2, \ldots, \tag{30.84}$$

где, как и прежде, $\theta_0 \equiv 1$. Константа $c(\theta)$ теперь не нужна, так как

$$\int_0^1 \left\{ \sum_{r=0}^{k} \theta_r \pi_r(y) \right\} dy = 1 + \sum_{r=1}^{k} \theta_r \int_0^1 \pi_r(y)\, dy = 1, \tag{30.85}$$

поскольку $\pi_0(y) \equiv 1$ и π_r ортогональны. Однако теперь надо обеспечить, чтобы (30.84) было неотрицательно на интервале (0, 1), и это накладывает ограничение на возможные значения θ_r. Таким образом, например, при $k = 1$ выражение для π_1, да-

ваемое (30.75), показывает, что мы должны ограничить θ_1 условием $|\theta_1| \leqslant 3^{-1/2}$.

Теперь, если мы положим $\theta_r = n^{-1/2}\lambda_r$, $r \geqslant 1$, то получим множество альтернатив, приближающихся к H_0 при $n \to \infty$. Более того, при $n \to \infty$ мы имеем

$$1 + n^{-1/2} \sum_r \lambda_r \pi_r(y) \sim \exp\left\{ n^{-1/2} \sum_r \lambda_r \pi_r(y) \right\},$$

так что асимптотическое распределение p_k^2 при альтернативах (30.76) будет применимо для (30.84) с $\theta_r = n^{-1/2}\lambda_r$. Для того чтобы получить асимптотическое нецентральное χ^2-распределение статистики p_k^2, в котором параметр нецентральности равен теперь $\lambda = \sum_{r=1}^{k} \lambda_r^2$, нам пришлось заставить H_k стремиться к H_0 при $n \to \infty$. Это в точности то же самое, что мы делали в **30.27**, чтобы получить соответствующий результат для критерия X^2.

30.42 Если мы имеем в виду конкретное альтернативное распределение $g(y)$, мы можем представить его через распределение класса (30.84) следующим образом. Выберем параметры θ_r в (30.84) так, чтобы минимизировать интеграл

$$Q^2 = \int_0^1 [g(y) - f(y \mid H_k)]^2 \, dy = \int_0^1 \left[g(y) - \left\{ 1 + \sum_{r=1}^{k} \theta_r \pi_r(y) \right\} \right]^2 dy. \tag{30.86}$$

Дифференцируя по θ_r, находим необходимые условия минимума

$$\int_0^1 \pi_r(y) \left[g(y) - \left\{ 1 + \sum_{r=1}^{k} \theta_r \pi_r(y) \right\} \right] dy = 0 \quad \text{при всех } r,$$

и, пользуясь ортогональностью $\pi_r(y)$, получаем отсюда

$$\mathbf{M}\{\pi_r(y) \mid H_k\} = \theta_r \quad \text{при всех } r. \tag{30.87}$$

Пользуясь еще раз ортогональностью, получаем минимальное значение (30.86) равным, как в обычной теории наименьших квадратов,

$$Q^2_{\min} = \int_0^1 \{g(y) - 1\}^2 \, dy - \int_0^1 \left\{ \sum_{r=1}^{k} \theta_r \pi_r(y) \right\}^2 dy =$$

$$= \int_0^1 g^2(y) \, dy - 1 - \sum_{r=1}^{k} \theta_r^2. \tag{30.88}$$

Величина $Q^2_{\min}$ неотрицательна по определению. Рассматриваемая как функция от k, она не возрастает, и так как от k зависит только $\sum_{r=1}^{k} \theta_r^2$, то $Q^2_{\min} \to 0$ при $k \to \infty$. Следовательно, при подборе модели (30.84) для $g(y)$ мы должны основываться на степени приближения параметра $\lambda = \sum_{r=1}^{k} \theta_r^2$ к своей верхней границе $\int_0^1 g^2(y)\,dy - 1$. Этот интеграл выражен в терминах величины, полученной в результате вероятностного интегрального преобразования, и часто бывает удобнее оценивать его в терминах альтернативного распределения первоначальной величины x. Назовем его $h(x)$. Мы имеем тогда, поскольку $g(y)dy = h(x)dx$,

$$\int_0^1 g^2(y)\,dy = \int_{-\infty}^{\infty} h^2(x)\frac{dx}{dy}\,dx = \int_{-\infty}^{\infty} \left\{ \frac{h^2(x)}{h(x \mid H_0)} \right\} dx. \qquad (30.89)$$

Пример 30.5

Рассмотрим нормальное распределение

$$h(x \mid \mu) = (2\pi)^{-1/2} \exp\left\{ -\tfrac{1}{2}(x-\mu)^2 \right\}, \quad -\infty \leqslant x \leqslant \infty,$$

с H_0: $\mu = 0$. Пользуясь (30.89), имеем

$$\int_0^1 g^2(y)\,dy = \int_{-\infty}^{\infty} \left\{ \frac{h^2(x \mid \mu)}{h(x \mid 0)} \right\} dx =$$

$$= (2\pi)^{-1/2} \int_{-\infty}^{\infty} \exp\left\{ -(x-\mu)^2 + \tfrac{1}{2}x^2 \right\} dx =$$

$$= \exp(\mu^2) \cdot (2\pi)^{-1/2} \int_{-\infty}^{\infty} \exp\left\{ -\tfrac{1}{2}(x-2\mu)^2 \right\} dx = \exp(\mu^2).$$

Таким образом, мы должны сравнивать $\lambda = \sum_{r=1}^{k} \theta_r^2$ с

$$\exp(\mu^2) - 1 = \mu^2 + \frac{\mu^4}{2!} + \ldots \qquad (30.90)$$

Из (30.87)

$$\theta_r = \int_0^1 \pi_r(y)\, g(y)\, dy.$$

Вследствие того, что $\pi_r\left(y-\frac{1}{2}\right)$ при четных r являются четными функциями (см. (30.75)) и $g(y)$ также четна относительно значения 1/2 при нашей симметричной альтернативе, мы имеем

$$\theta_{2r}=0.$$

При нечетных r мы должны вычислять отдельно каждый член. Пользуясь (30.75), находим

$$\theta_1=\int_{-1/2}^{1/2}3^{1/2}\cdot 2zg(z)\,dz=3^{1/2}\cdot 2\int_{-\infty}^{\infty}zh(x)\,dx=$$

$$=3^{1/2}\cdot 2\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{x}h(u\mid H_0)\,du-\frac{1}{2}\right\}h(x\mid H_0)\exp\left\{-\frac{1}{2}(\mu^2-2\mu x)\right\}dx$$

и

$$\left[\frac{d\theta_1}{d\mu}\right]_{\mu=0}=3^{1/2}\cdot 2\int_{-\infty}^{\infty}x\left\{\int_{-\infty}^{x}h(u\mid H_0)\,du-\frac{1}{2}\right\}h(x\mid H_0)\,dx=3^{1/2}\cdot\frac{1}{2}\,\Delta_1,$$

где Δ_1 — средняя разность Джини (см. упражнение 2.9), равная $2/\pi^{1/2}$ в нормальном случае (см. **10.14**), так что

$$\left[\frac{d\theta_1}{d\mu}\right]=\left(\frac{3}{\pi}\right)^{1/2}.$$

Таким образом, при малом отклонении $d\mu$ параметра μ величина θ_1 изменится на $(3/\pi)^{1/2}\,d\mu\approx 0{,}98\,d\mu$, и, если мы применяем критерий p_1^2 с $\theta_1=(3/\pi)^{1/2}\mu$, мы видим из формулы (30.90), что $\theta_1^2\approx(3/\pi)\mu^2$ будет лишь немногим меньше ее правой части, и потеря эффективности при проверке отклонения среднего μ от нуля в нормальном распределении будет мала. Это легко подтвердить следующим образом. Согласно **30.41** p_1^2 имеет асимптотически нецентральное распределение χ^2 с 1 степенью свободы и параметром нецентральности $n(3/\pi)\mu^2$, эквивалентным нормированному нормальному отклонению $(3n/\pi)^{1/2}\mu$. Наилучший критерий, основанный на выборке, использует нормальное отклонение $n^{1/2}\mu$. Множитель

$$(3/\pi)^{1/2}=0{,}98$$

мало влияет на мощность при больших выборках.

Преимущество критериев p_k^2, продемонстрированное здесь, состоит в том, что при заданной альтернативе мы можем выбирать k так, чтобы получить более высокую мощность, чем с критерием X^2.

30.43 Начиная с **30.37**, мы ограничивались простой гипотезой и негруппированными наблюдениями. Теперь мы перейдем к обсуждению группировки данных в критериях p_k^2 (что очень нужно для практики ввиду необходимости производить вероятностное интегральное преобразование над каждым наблюдением) и распространения критериев на сложные гипотезы, что, может быть, еще более важно. Эти вопросы составляют основное содержание второй и третьей работ Бартона (1953, 1955, 1956). Замечателен тот факт, что, как только произведено группирование, обнаруживается близкая взаимосвязь между критериями p_k^2 и критерием X^2.

Предположим, что область значений варианты x разбита на k классов, и пусть ξ_i — медиана i-й по величине группы, считая снизу. Очевидным аналогом (30.73) тогда служит

$$y_i' = \int_{-\infty}^{\xi_i} f_0(u)\,du = \sum_{j=1}^{i-1} p_{0j} + \frac{1}{2} p_{0i}, \tag{30.91}$$

где p_{0i}, как и раньше, — гипотетические вероятности. Мы заменяем все y_i одного класса значением y_i' и пишем, как раньше $z_i' = y_i' - 1/2$. Нам теперь требуется набор ортогональных полиномов $P_r(y')$, которые будут играть ту же роль для группированных данных, что и нормированные полиномы Лежандра для негруппированных. Ввиду того, что альтернативная гипотеза может быть теперь сформулирована в терминах переменной x как

$$f(x \mid H_s) = \left\{ \sum_{r=0}^{s} \theta_r P_r(y') \right\} f(x \mid H_0),$$

мы после группировки по k классам получаем выражение для вероятностей при альтернативе через гипотетические вероятности:

$$p_{1i} \mid H_s = \left\{ \sum_{r=0}^{s} \theta_r P_r(y') \right\} p_{0i}. \tag{30.92}$$

Естественно поэтому задать $P_r(y')$ соотношениями

$$P_0(y') \equiv 1, \tag{30.93}$$

$$\sum_{i=1}^{k} p_{0i} P_r(y_i') P_t(y_i') = \begin{cases} 1, & r = t, \\ 0, & r \neq t, \ r, \ t = 0, 1, 2, \ldots, k-1. \end{cases} \tag{30.94}$$

Тогда (30.80) (с заменой k на $k-1$) вследствие (30.91) принимает вид

$$p_{k-1}^2 = \sum_{r=1}^{k-1} u_r^2 = \frac{1}{n} \sum_{r=1}^{k-1} \left\{ \sum_{i=1}^{n} P_r(y_i) \right\}^2 = \frac{1}{n} \sum_{r=1}^{k-1} \left\{ \sum_{i=1}^{k} n_i P_r(y_i') \right\}^2,$$

что ввиду (30.93) равно

$$\frac{1}{n}\sum_{r=0}^{k-1}\left\{\sum_{i=1}^{k} n_i P_r(y_i')\right\}^2 - n = \sum_{r=0}^{k-1}\left\{\sum_{i=1}^{k}\frac{n_i}{(np_{0i})^{1/2}}\cdot p_{0i}^{1/2}P_r(y_i')\right\}^2 - n.$$

Здесь суммируются квадраты взвешенных сумм величин $p_{0i}^{1/2}P_r(y_i')$ с весами $n_i/(np_{0i})^{1/2}$. Вследствие свойства ортогональности (30.94) мы имеем

$$p_{k-1}^2 + n = \sum_{i=1}^{k}\frac{n_i^2}{np_{0i}},$$

или, ввиду (30.8), в точности

$$p_{k-1}^2 = X^2. \tag{30.95}$$

30.44 Так же как p_{k-1}^2 совпадает вследствие (30.95) с X^2, критерии низших порядков p_r^2 $(r = 1, 2, \ldots, k-2)$, как мы теперь увидим, являются компонентами или разбиениями X^2, т. е. определенного вида функциями от асимптотически нормированных нормальных величин

$$x_i = \frac{n_i - np_{0i}}{(np_{0i})^{1/2}},$$

которые мы теперь рассмотрим. Если обозначить

$$u_r = \sum_{i=1}^{k} l_{ri}x_i, \qquad r = 1, 2, \ldots, k, \tag{30.96}$$

и последним столбцом матрицы $\boldsymbol{L} = \{l_{ri}\}$ взять

$$l_{ki} = p_{0i}^{1/2},$$

то получим тождественно

$$u_k = 0,$$

а если выбрать остальные элементы матрицы $\boldsymbol{L}$ так, чтобы она была ортогональна, т. е.

$$\sum_{i=1}^{k} l_{ri}l_{si} = \begin{cases} 1, & r = s, \\ 0, & r \neq s, \end{cases}$$

то u_r также будут асимптотически нормированными нормальными величинами и, вследствие ортогональности, асимптотически независимыми. Поэтому сумма квадратов любых m $(m = 1, 2, \ldots, k-1)$ из u_r $(r = 1, 2, \ldots, k-1)$ будет распределена так же, как χ^2 с m степенями свободы, и не будет зависеть от любой суммы, основанной на других u_r. Мы вернемся к задачам, связанным с разбиениями X^2, в определенном контексте главы 33.

С этой точки зрения достоинство критериев p_1^2, p_2^2, ... с группированными данными состоит в подборе подходящих функций от y_i для получения критерия с наибольшей мощностью; они выделяют важнейшие компоненты X^2.

30.45 Имея в виду замечания в **30.44**, мы не будем удивлены тем, что при переходе к сложным гипотезам теория группированных критериев p_k^2 близко напоминает теорию критерия X^2, которая уже обсуждалась в **30.11—21**. Как показал Бартон (см. Ватсон (1959)), переносятся все принципиальные результаты: если используются мультиномиальные МП-оценки, то число степеней свободы уменьшается на количество оцениваемых параметров; если способ группирования определяется по наблюдениям, то это не вносит изменений (при условиях регулярности) в асимптотические распределения.

Главная проблема при применении критериев p_k^2 к сложным гипотезам — это выбор k. Как мы отмечали в случае простой гипотезы, при проверке согласия часто не имеют в виду определенной альтернативы, — в противном случае лучше было бы применить по возможности более специфический критерий. Ввиду того, что большие выборки часто используются при проверке согласия, так что группирование наблюдений является практической необходимостью, совпадение группированного критерия p_{k-1}^2 с критерием X^2 означает, что помимо задач о разбиениях, общих для обоих типов критериев, между этими критериями нет конкуренции.

Критерии согласия, основанные на выборочной функции распределения

30.46 Все остальные распространенные критерии согласия являются функциями кумулятивного распределения выборки, или *выборочной функции распределения*, определяемой как

$$S_n(x) = \begin{cases} 0, & x < x_{(1)}, \\ \dfrac{r}{n}, & x_{(r)} \leqslant x < x_{(r+1)}, \\ 1, & x_{(n)} \leqslant x. \end{cases} \tag{30.97}$$

Здесь $x_{(r)}$ — порядковые статистики, т. е. наблюдения, упорядоченные так, что

$$x_{(1)} \leqslant x_{(2)} \leqslant \ldots \leqslant x_{(n)}.$$

$S_n(x)$ — это просто относительное число наблюдений, не превосходящих x. Если $F_0(x)$ — полностью определенная истинная ф. р.,

из которой получены наблюдения, мы имеем для каждого значения x согласно усиленному закону больших чисел

$$P\{\lim_{n\to\infty} S_n(x) = F_0(x)\} = 1, \tag{30.98}$$

и на самом деле имеются и более сильные результаты относительно сходимости выборочной ф. р. к истинной ф. р.

В определенном смысле (30.98) — это фундаментальное соотношение, на котором основана вся статистическая теория. Если бы не выполнялось никакого соотношения такого рода, то случайная выборка не имела бы смысла. В нашем теперешнем контексте понятно, что критерий согласия может основываться на любой мере расхождения между $S_n(x)$ и $F_0(x)$. Мы теперь предположим, что $F_0(x)$ непрерывна. Рассмотрим статистику

$$W^2 = \int_{-\infty}^{\infty} \{S_n(x) - F_0(x)\}^2 \, dF_0(x), \tag{30.99}$$

которая была предложена Смирновым (1936) после более ранних предложений Крамера и Мизеса. Из теории биномиального распределения (пример 3.2) с $p = F_0(x)$ следует, что

$$\mathbf{M}\{S_n(x) - F_0(x)\}^2 = F_0(x)\{1 - F_0(x)\}/n. \tag{30.100}$$

Поэтому из (30.99) и (30.100) мы имеем

$$\mathbf{M}(W^2) = \frac{1}{n}\int_0^1 F_0(1 - F_0)\, dF_0 = \frac{1}{n}\left(\frac{1}{2} - \frac{1}{3}\right) = \frac{1}{6n}, \tag{30.101}$$

и аналогично можно установить, что

$$\mathbf{D}(W^2) = \mathbf{M}(W^4) - \mathbf{M}^2(W^2) = \frac{4n-3}{180n^3}. \tag{30.102}$$

30.47 Можно заметить, что среднее и дисперсия статистики W^2 не зависят от F_0. На самом деле распределение W^2 в целом не зависит от F_0: этот критерий полностью свободен от распределения при любом n. Это легко увидеть непосредственно, поскольку если мы применим вероятностное интегральное преобразование (30.73) к x, то мы сведем (30.99) к

$$W^2 = \int_0^1 \{S_n(y) - y\}^2 \, dy, \tag{30.103}$$

т. е. задача проверки согласия сводится к проверке того, насколько велико в выборке из равномерного распределения на (0, 1) отклонение выборочной ф. р. от ф. р. этого распределения, $F(y) = y$.

Из (30.101) и (30.102) ясно, что при отыскании предельного распределения следует рассматривать nW^2 (нормирующий множитель здесь не равен $n^{1/2}$, как обычно требуется по центральной предельной теореме), тогда среднее и дисперсия имеют нулевой порядок по n. Асимптотическая теория nW^2 трудна, а точной теории для конечных n неизвестно. Смирнов (1936) показал, что ее предельная х. ф. равна

$$\varphi(t) = \lim_{n\to\infty} \mathbf{M}\{\exp(itnW^2)\} = \left\{\frac{(2it)^{1/2}}{\sin[(2it)^{1/2}]}\right\}^{1/2}. \qquad (30.104)$$

Андерсон и Дарлинг (1952) преобразовали $\varphi(t)$ к виду, пригодному для численных расчетов, и табулировали обратную функцию для предельной ф. р. nW^2. В таблице даны значения, которые nW^2 превышает с вероятностями 0,001; 0,01(0,01)0,99 *). Как обычно принято, наиболее важными из этих значений для целей проверки гипотез являются:

Размер критерия α	Критическое значение nW^2
0,10	0,347
0 05	0,461
0,01	0,743
0,001	1,168

Критическую область образуют большие значения nW^2, что очевидно из мотивировки критерия.

Маршалл (1958) показал, что асимптотическое распределение nW^2 достигается чрезвычайно быстро, так что приведенные выше критические значения достаточно точны уже при $n = 3$ Marshall A. W. (1958), The small-sample distribution of $n\omega_n^2$ Ann Math. Statist. **29**, 307).

Пирсон и Стефенс (1962) применили подгонку распределений Джонсона типа S_B (см. **6.27—34**) при $n = 5$, 10, ∞ для получения критических значений.

30.48 Для критерия W^2, так же как для негруппированных критериев p_k^2, обсуждавшихся в **30.37—42**, требуется вычислять $F_0(x)$ для каждого отдельного наблюдения. Действительно, можно показать, что эта статистика может быть представлена в виде

$$nW^2 = \frac{1}{12n} + \sum_{r=1}^{n}\left\{F_0(x_{(r)}) - \frac{2r-1}{2n}\right\}^2. \qquad (30.105)$$

*) Эта таблица воспроизведена в *Сборнике статистических таблиц*. Пятизначные таблицы самой предельной ф. р. nW^2, а также предельной ф. р. статистики Андерсона — Дарлинга (см. ниже, в **30.48**) имеются в *Таблицах математической статистики*. (*Прим. ред.*)

Критерий W^2 рассматривался в случае сложной гипотезы с одним неизвестным параметром Дарлингом (1955). Тогда статистика критерия, вообще говоря, больше не является свободной от распределения, как и следовало ожидать из рассмотрений в **30.36**. Исключение составляет случай, когда параметр допускает оценку с дисперсией меньшего порядка, чем n^{-1}; тогда предельное распределение оказывается точно таким же, как и при простой гипотезе (ср. **30.8**, где мы встретились с тем же явлением для X^2). Если неизвестен параметр сдвига или масштаба, оцениваемый с дисперсией порядка n^{-1}, то предельное распределение будет свободным от параметра (ср. **30.36**).

Андерсон и Дарлинг (1952, 1954) исследовали еще одну статистику критерия для простой гипотезы. Она дается формулой (30.99) с добавлением множителя $[F_0(x)\{1-F_0(x)\}]^{-1}$ под интегралом. Они табулировали критические значения ее асимптотического распределения для $\alpha = 0{,}10;\ 0{,}05;\ 0{,}01$. П. Льюис (1961) дает точную таблицу этой ф. р. для $n = 1$ и $n \to \infty$ и оценки, основанные на выборочных экспериментах для $n = 2(1)8$; сходимость с асимптотической ф. р. чрезвычайно быстрая.

Ватсон (1961) показал, что если видоизменить статистику nW^2, полагая

$$U^2 = n\int_0^1 \left[S_n(x) - F_0(x) - \int_0^1 \{S_n(x) - F_0(x)\}\, dF_0(x)\right]^2 dF_0(x),$$

то $\pi^2 U^2$ будет иметь точно такое же асимптотическое распределение, как nD_n^2, даваемое в (30.132) ниже. Пирсон и Стефенс (1962) и Стефенс (1963, 1964) дают теоретические и эмпирические результаты о распределении U^2. Тайкью (1965b) использует распределение χ^2 для приближения распределений U^2, а также W^2.

Статистика Колмогорова

30.49 Мы теперь переходим к наиболее важному из общих критериев согласия, отличных от X^2. Подобно статистике W^2, определенной в (30.99), он основан на отклонениях выборочной ф. р. $S_n(x)$ от полностью заданной непрерывной гипотетической ф. р. $F_0(x)$. Однако здесь используется более простая мера отклонения — максимум абсолютной величины разности между $S_n(x)$ и $F_0(x)$. Мы, таким образом, определим

$$D_n = \sup_x |S_n(x) - F_0(x)|. \tag{30.106}$$

Из-за наличия модуля в определении (30.106) мы могли бы ожидать трудностей в исследовании распределения D_n, но замечательно, что асимптотическое распределение было получено Колмогоровым (1933), когда он впервые предложил эту статистику. Нижеследующий вывод принадлежит Феллеру (1948).

30.50 Мы прежде всего отметим, что распределение D_n полностью свободно от распределения, когда выполнена H_0. В этом случае это можно видеть непосредственно, так как если начертить $S_n(x)$ и $F_0(x)$ как ординаты, соответствующие абсциссе x, то D_n будет просто значением наибольшего расхождения между ними по вертикали. Ясно, что если мы сделаем любое взаимно

однозначное преобразование x, то оно не изменит расстояния по вертикали в любой точке и, в частности, значение D_n останется без изменения.

30.51 Теперь рассмотрим значения $x_{10}, x_{20}, \ldots, x_{n-1,0}$, определяемые равенством

$$F_0(x_{k0}) = k/n. \tag{30.107}$$

(Если при каком-нибудь k (30.107) выполняется в пределах некоторого интервала, то мы берем в качестве x_{k0} нижнюю границу этого интервала.) Пусть c — положительное число. Если при некотором значении x

$$S_n(x) - F_0(x) > c/n, \tag{30.108}$$

то неравенство (30.108) будет выполняться для всех x из некоторого интервала, в правой граничной точке которого, x', оно обращается в равенство, т. е.

$$S_n(x') - F_0(x') = c/n. \tag{30.109}$$

Так как по определению $S_n(x)$ — ступенчатая функция, принимающая значения, кратные $1/n$, а c — целое число, то из (30.109) следует, что $F_0(x')$ кратно $1/n$, и, таким образом, вследствие (30.107) $x' = x_{k0}$ для некоторого k, так что (30.109) принимает вид

$$S_n(x_{k0}) - F_0(x_{k0}) = c/n,$$

т. е. из (30.107)

$$S_n(x_{k0}) = (k + c)/n. \tag{30.110}$$

По определению $S_n(x)$ в (30.97) это означает, что в точности $k + c$ наблюденных значений x меньше x_{k0}, являющегося гипотетическим значением, ниже которого должны были бы попасть k наблюдений. Обратно, если $x_{(k+c)} \leqslant x_{k0} < x_{(k+c+1)}$, то отсюда сразу вытекает (30.108). Мы, следовательно, установили предварительный результат, что неравенство

$$S_n(x) - F_0(x) > c/n$$

выполняется при некотором x тогда и только тогда, когда при каком-нибудь k

$$x_{(k+c)} \leqslant x_{k0} < x_{(k+c+1)}. \tag{30.111}$$

Следовательно, мы можем ограничиться рассмотрением вероятности того, что выполняется (30.111).

30.52 Мы обозначим событие (30.111) через $A_k(c)$. Из (30.106) мы видим, что статистика D_n будет превосходить c/n тогда и только тогда, когда происходит по крайней мере одно из $2n$ событий

$$A_1(c),\ A_1(-c),\ A_2(c),\ A_2(-c),\ \ldots,\ A_n(c),\ A_n(-c). \tag{30.112}$$

Мы определим теперь $2n$ взаимно несовместных событий U_r и V_r. Событие U_r происходит, если первым в последовательности (30.112) осуществляется событие $A_r(c)$, а V_r происходит, если первым осуществляется $A_r(-c)$. Очевидно,

$$\mathsf{P}\left\{D_n > \frac{c}{n}\right\} = \sum_{r=1}^{n} [\mathsf{P}\{U_r\} + \mathsf{P}\{V_r\}]. \tag{30.113}$$

Из определений $A_k(c)$ и U_r, V_r имеем соотношения

$$\left.\begin{aligned} \mathsf{P}\{A_k(c)\} &= \sum_{r=1}^{k} [\mathsf{P}\{U_r\}\mathsf{P}\{A_k(c) \mid A_r(c)\} + \\ &\qquad + \mathsf{P}\{V_r\}\mathsf{P}\{A_k(c) \mid A_r(-c)\}], \\ \mathsf{P}\{A_k(-c)\} &= \sum_{r=1}^{k} [\mathsf{P}\{U_r\}\mathsf{P}\{A_k(-c) \mid A_r(c)\} + \\ &\qquad + \mathsf{P}\{V_r\}\mathsf{P}\{A_k(-c) \mid A_r(-c)\}], \end{aligned}\right\} \tag{30.114}$$

Из (30.111) и (30.107) мы видим, что $\mathsf{P}\{A_k(c)\}$ есть вероятность того, что в n биномиальных испытаниях с вероятностью k/n произойдут в точности $k + c$ «успехов», т. е.

$$\mathsf{P}\{A_k(c)\} = \binom{n}{k+c}\left(\frac{k}{n}\right)^{k+c}\left(1 - \frac{k}{n}\right)^{n-(k+c)}. \tag{30.115}$$

Аналогично, при $r \leqslant k$

$$\left.\begin{aligned} \mathsf{P}\{A_k(c) \mid A_r(c)\} &= \binom{n-(r+c)}{k-r}\left(\frac{k-r}{n-r}\right)^{k-r}\left(1 - \frac{k-r}{n-r}\right)^{n-(k+c)}, \\ \mathsf{P}\{A_k(c) \mid A_r(-c)\} &= \binom{n-(r-c)}{k-r+2c}\left(\frac{k-r}{n-r}\right)^{k-r+2c}\left(1 - \frac{k-r}{n-r}\right)^{n-(k+c)}. \end{aligned}\right\} \tag{30.116}$$

Соотношения (30.115) и (30.116) выполняются как для положительных, так и для отрицательных c. Применяя их, мы видим, что (30.114) представляет собой систему из $2n$ линейных уравнений от $2n$ неизвестных $\mathsf{P}\{U_r\}$, $\mathsf{P}\{V_r\}$. Решив их и подставив решение в (30.113), мы получили бы $\mathsf{P}\left\{D_n > \frac{c}{n}\right\}$ для любого c.

30.53 Если мы теперь обозначим

$$p_k(c) = e^{-k}\frac{k^{k+c}}{(k+c)!}, \tag{30.117}$$

то

$$\left.\begin{aligned} \mathsf{P}\{A_k(c)\} &= p_k(c)\, p_{n-k}(-c)/p_n(0), \\ \mathsf{P}\{A_k(c) \mid A_r(c)\} &= p_{k-r}(0)\, p_{n-k}(-c)/p_{n-r}(-c), \\ \mathsf{P}\{A_k(c) \mid A_r(-c)\} &= p_{k-r}(2c)\, p_{n-k}(-c)/p_{n-r}(c), \end{aligned}\right\} \tag{30.118}$$

так что если определить

$$u_r = \mathsf{P}\{U_r\}\frac{p_n(0)}{p_{n-r}(-c)}, \qquad v_r = \mathsf{P}(V_r)\frac{p_n(0)}{p_{n-r}(c)} \tag{30.119}$$

и подставить (30.115)—(30.119) в (30.114), то это соотношение принимает вид

$$\left.\begin{aligned} p_k(c) &= \sum_{r=1}^{k}[u_r p_{k-r}(0) + v_r p_{k-r}(2c)], \\ p_k(-c) &= \sum_{r=1}^{k}[u_r p_{k-r}(-2c) + v_r p_{k-r}(0)]. \end{aligned}\right\} \tag{30.120}$$

Из системы (30.120) нам нужно найти

$$\sum_{r=1}^{n}[\mathsf{P}\{U_r\} + \mathsf{P}\{V_r\}] = \frac{1}{p_n(0)}\sum_{r=1}^{n}[p_{n-r}(-c)\,u_r + p_{n-r}(c)\,v_r]. \tag{30.121}$$

Мы поэтому определим

$$p_k = \frac{1}{p_n(0)}\sum_{r=1}^{k} p_{k-r}(-c)\,u_r, \qquad q_k = \frac{1}{p_n(0)}\sum_{r=1}^{k} p_{k-r}(c)\,v_r, \tag{30.122}$$

так что из (30.121)

$$\sum_{r=1}^{k}[\mathsf{P}\{U_r\} + \mathsf{P}\{V_r\}] = p_n + q_n. \tag{30.123}$$

Теперь мы введем производящие функции для p_k и q_k, а именно:

$$G_p(t) = \sum_{k=1}^{\infty} p_k t^k, \qquad G_q(t) = \sum_{k=1}^{\infty} q_k t^k.$$

Если мы определим также производящие функции для u_k, v_k и (для удобства) $n^{-1/2}p_k(c)$, а именно:

$$G_u(t) = \sum_{k=1}^{\infty} u_k t^k, \qquad G_v(t) = \sum_{k=1}^{\infty} v_k t^k$$

и

$$G(t,\ c) = n^{-1/2}\sum_{k=1}^{\infty} p_k(c)\,t^k,$$

то из (30.122) получим соотношения

$$\left.\begin{aligned} G_p(t) &= G_u(t)\,G(t,\ -c)\,n^{1/2}/p_n(0), \\ G_q(t) &= G_v(t)\,G(t,\ c)\,n^{1/2}/p_n(0). \end{aligned}\right\} \tag{30.124}$$

30.54 Мы рассмотрим теперь предельную форму (30.124). Положим

$$c = zn^{1/2},$$

и пусть $n \to \infty$ и $c \to \infty$ так, чтобы z оставалось фиксированным.

Мы видим из (30.117), что $p_k(c)$ — это просто вероятность значения $(k + c)$ для пуассоновской случайной величины с параметром k, т. е. вероятность того, что она превосходит свое среднее значение на $c/k^{1/2}$ стандартных отклонений. Если k/n стремится к некоторому фиксированному значению m, то, поскольку пуассоновское распределение сходится к нормальному,

$$p_k(c) \to (2\pi k)^{-1/2} \exp\left(-\frac{1}{2}\frac{c^2}{k}\right);$$

полагая $k = mn$, $c = zn^{1/2}$, имеем

$$n^{1/2} p_k(zn^{1/2}) \to (2\pi m)^{-1/2} \exp\left(-\frac{1}{2}\frac{z^2}{m}\right). \qquad (30.125)$$

Так как $G(t, c)$ — производящая функция для $n^{-1/2}p_k(c)$, то

$$G(e^{-t/n}, zn^{1/2}) = n^{-1/2} \sum_{k=1}^{\infty} p_k(zn^{1/2})\, e^{-tk/n},$$

и при нашем предельном переходе это выражение в силу (30.125) стремится к

$$\lim_{n\to\infty} G(e^{-t/n}, zn^{1/2}) = (2\pi)^{-1/2} \int_0^{\infty} m^{-1/2} \exp\left(-tm - \frac{1}{2}\frac{z^2}{m}\right) dm. \qquad (30.126)$$

Если мы продифференцируем интеграл I в правой части (30.126) по $z^2/2$, то получим простое дифференциальное уравнение

$$\frac{\partial I}{\partial (z^2/2)} = -\left(\frac{t}{z^2/2}\right)^{1/2} I,$$

решением которого является

$$I = \left(\frac{\pi}{t}\right)^{1/2} \exp\{-(2tz^2)^{1/2}\}.$$

Таким образом,

$$\lim_{n\to\infty} G(e^{-t/n}, zn^{1/2}) = (2t)^{-1/2} \exp\{-(2tz^2)^{1/2}\}. \qquad (30.127)$$

Выражение (30.127) представляет собой четную функцию от z, а следовательно, и от c.

Так как вследствие (30.120)

$$\left.\begin{aligned} G(t, c) &= G_u(t)\, G(t, 0) + G_v(t)\, G(t, 2c), \\ G(t, -c) &= G_u(t, -2c) + G_v(t)\, G(t, 0), \end{aligned}\right\} \qquad (30.128)$$

четность (30.127) по c дает нам

$$\lim_{n\to\infty} G_u(e^{-t/n}) = \lim_{n\to\infty} G_v(e^{-t/n}) =$$

$$= \frac{\lim G(e^{-t/n}, zn^{1/2})}{\lim G(e^{-t/n}, 0) + \lim (e^{-t/n}, 2zn^{1/2})} = \frac{\exp\{-(2tz^2)^{1/2}\}}{1+\exp\{-(8tz^2)^{1/2}\}} \qquad (30.129)$$

согласно (30.127). Поэтому, применяя (30.127) и (30.129) в соотношениях (30.124) и вспоминая, что

$$p_n(0) \sim (2\pi n)^{-1/2},$$

получаем

$$\lim_{n\to\infty} n^{-1}G_p(e^{-t/n}) = \lim_{n\to\infty} n^{-1}G_q(e^{-t/n}) =$$

$$= \left(\frac{2\pi}{2t}\right)^{1/2} \frac{\exp\{-(8tz^2)^{1/2}\}}{1+\exp\{-(8tz^2)^{1/2}\}} = L(t).$$

Это можно разложить в геометрический ряд

$$L(t) = \left(\frac{2\pi}{2t}\right)^{1/2} \sum_{r=1}^{\infty} (-1)^{r-1} \exp\{-(8tr^2z^2)^{1/2}\}. \qquad (30.130)$$

С помощью такого же интегрирования, как в (30.126), можно видеть, что $L(t)$ есть одностороннее преобразование Лапласа $\int_0^\infty e^{-mt} f(m)\, dm$ от функции

$$f(m) = \sum_{r=1}^{\infty} (-1)^{r-1} \exp\{-2r^2z^2/m\}. \qquad (30.131)$$

Таким образом, (30.131) является результатом обращения любой из предельных производящих функций p_k или q_k, первая из которых есть

$$\lim n^{-1}G_p(e^{-t/n}) = \lim n^{-1} \sum_{k=1}^{\infty} p_k e^{-tk/n} = \int_0^\infty (\lim p_k)\, e^{-tm}\, dm.$$

В силу (30.113) и (30.123) нам требуется только значение $(p_n + q_n)$. Поэтому мы полагаем в (30.131) $k = n$, т. е. $m = 1$, и после умножения на два получаем наш окончательный результат

$$\lim_{n\to\infty} \mathsf{P}\{D_n > zn^{-1/2}\} = 2 \sum_{r=1}^{\infty} (-1)^{r-1} \exp\{-2r^2z^2\}. \qquad (30.132)$$

Смирнов (1948) табулировал функцию (30.132) (на самом деле ее дополнение) для $z = 0{,}28(0{,}01)2{,}50(0{,}05)3{,}00$ с шестью десятичными знаками

или более. Это полностью покрывает эффективную область предельного распределения.

30.55 Помимо вывода предельного результата (30.132) Колмогоров (1933) дал рекуррентные соотношения для конечных n, которые затем были использованы для табулирования распределения D_n. Бирнбаум (1952) дает таблицы $\mathsf{P}\{D_n < c/n\}$ с 5 десятичными знаками для $n = 1(1)100$ и $c = 1(1)15$ и обратные таблицы значений D_n, для которых эта вероятность равна 0,95, для $n = 2(1)5(5)30(10)100$ и значений D_n, для которых эта вероятность равна 0,99, для $n = 2(1)5(5)30(10)80$. Л. Миллер (1956) дает обратные таблицы *) для $n = 1(1)100$ и вероятностей 0,90, 0,95, 0,98, 0,99. До этого Мэсси (1950a, 1951a) дал $\mathsf{P}\{D_n < c/n\}$ для $n = 5(5)80$ и некоторых значений $c \leqslant 9$, а также обратные таблицы для $n = 1(1)20(5)35$ и вероятностей 0,80, 0,85, 0,90, 0,95, 0,99.

Критические значения асимптотического распределения оказываются равными:

Размер критерия	Критическое значение D_n
0,05	$1{,}3581n^{-1/2}$
0,01	$1{,}6276n^{-1/2}$

причем эти значения всегда больше, чем точные значения при конечных n. Точность приближения для этих значений α удовлетворительна при $n = 80$.

Доверительные границы для функции распределения

30.56 Поскольку распределение D_n свободно от распределения и достаточно хорошо известно для всех n и поскольку в качестве меры расхождения используется максимальное абсолютное отклонение $S_n(x)$ от $F(x)$, мы можем обратить процедуру проверки согласия и использовать D_n для установления доверительных границ для (непрерывной) функции распределения *в целом*. Действительно, *какова бы ни была истинная ф. р.* $F(x)$, мы имеем, обозначая d_α критическое значение D_n при размере критерия α,

$$\mathsf{P}\{D_n = \sup_x |S_n(x) - F(x)| > d_\alpha\} = \alpha.$$

Поэтому мы можем обратить это в доверительное утверждение

$$\mathsf{P}\{S_n(x) - d_\alpha \leqslant F(x) \leqslant S_n(x) + d_\alpha \text{ при всех } x\} = 1 - \alpha. \quad (30.133)$$

Таким образом, мы просто располагаем полосу шириной $\pm d_\alpha$ вокруг выборочной ф. р. $S_n(x)$, и с вероятностью $1 - \alpha$ истинная ф. р. $F(x)$ лежит целиком внутри этой полосы. Это замечательно простой и прямой метод оценки функции распределения. Никакой другой критерий согласия не позволяет такого обращения критерия в доверительный интервал, так как ни один критерий

*) Таблицы Смирнова ф. р. (30.132) и таблицы Миллера процентных точек приведены в *Таблицах математической статистики* и *Сборнике статистических таблиц*. В последнем имеются также таблицы Бирнбаума для $\mathsf{P}\{D_n < c/n\}$. (*Прим. ред.*)

не использует такой прямой и просто интерпретируемой меры расхождения, как D_n.

Из этой техники доверительных интервалов можно извлечь полезные заключения относительно объема выборки, необходимого для близкой аппроксимации функции распределения. Например, из критических значений, приведенных в конце **30.55**, следует, что эмпирическая функция распределения выборки в 100 наблюдений с вероятностью 0,95 отстояла бы повсюду не далее чем на 0,13581 от истинной ф. р. Для того чтобы она лежала повсюду в пределах 0,05 от истинной ф. р. с вероятностью 0,99, потребовался бы объем выборки $(1{,}6276/0{,}05)^2$, т. е. более чем 1000.

Нётер (1963) показал, что для дискретных распределений вероятность в левой части (30.133) $\geqslant 1-\alpha$. Таким образом, в этом случае критерий D_n также консервативен.

30.57 Будучи основана на абсолютном значении, величина D_n не позволяет строить односторонних доверительных интервалов для $F(x)$, но мы можем рассматривать только положительные отклонения и определить

$$D_n^+ = \sup_x \{S_n(x) - F_0(x)\}, \qquad (30.134)$$

как было сделано Вальдом и Волфовицем (1939) и Смирновым (1939а).

Чтобы получить предельное распределение D_n^+, мы заново проследим за рассуждениями в **30.51—54**. Мы теперь рассматриваем только события $A_k(c)$ с $c > 0$ в (30.112). U_r определяется, как и раньше, а V_r не рассматривается. (30.114) заменяется на

$$\mathbf{P}\{A_k(c)\} = \sum_{r=1}^{k} \mathbf{P}\{U_r\}\,\mathbf{P}\{A_k(c) \mid A_r(c)\},$$

а (30.128) — на

$$G(t,\ c) = G_u(t)\,G(t,\ 0). \qquad (30.135)$$

Вместо (30.129) мы, следовательно, имеем, используя (30.127) и (30.135),

$$\lim_{n\to\infty} G_u(e^{-t/n}) = \exp\{-(2tz^2)^{1/2}\}.$$

Первое уравнение в (30.124) справедливо, и мы получаем таким же способом, как и раньше,

$$\lim_{n\to\infty} n^{-1}G_p(e^{-t/n}) = \left(\frac{2\pi}{2t}\right)^{1/2} \exp\{-(8tz^2)^{1/2}\}. \qquad (30.136)$$

Снова из (30.127) видно, что (30.136) есть одностороннее преобразование Лапласа функции

$$f(m) = m^{-1/2}\exp(-2z^2/m),$$

и подстановка $m = 1$, как и раньше, дает

$$\lim_{n\to\infty} \mathbf{P}\{D_n^+ > zn^{-1/2}\} = \exp(-2z^2); \qquad (30.137)$$

это результат Смирнова (1939а). Соотношение (30.137) можно переписать в виде

$$\lim_{n\to\infty} \mathbf{P}\{2n(D_n^+)^2 \leqslant 2z^2\} = 1 - \exp(-2z^2). \qquad (30.138)$$

Дифференцирование (30.138) по $(2z^2)$ показывает, что величина $y = 2n(D_n^+)^2$ имеет асимптотически экспоненциальное распределение

$$dF(y) = \exp(-y)\,dy, \quad 0 \leqslant y \leqslant \infty.$$

Иначе говоря, $2y = 4n(D_n^+)^2$ имеет асимптотически распределение χ^2 с 2 степенями свободы. Очевидно, в точности та же теория будет справедлива, если мы рассмотрим только отрицательные отклонения.

30.58 Бирнбаум и Тинджи (1951) получили выражение для точного распределения величины D_n^+ и табулировали значения, которые она превышает с вероятностями 0,10, 0,05, 0,01, 0,001, для $n = 5, 8, 10, 20, 40, 50$. Как и для D_n, асимптотические значения превышают точные и разности между ними малы при $n = 50$ *).

Мы можем, очевидно, использовать D_n^+ для получения односторонних доверительных областей вида $\mathbf{P}\{S_n(x) - d_\alpha^+ \leqslant F(x)\} = 1 - \alpha$, где d_α^+ — критическое значение D_n^+.

Сравнение статистики Колмогорова с X^2

30.59 В общем случае ничего не известно о поведении статистики D_n, когда при проверке сложной гипотезы согласия подлежат оценке параметры, хотя применение этой статистики для проверки нормальности изучалось (см. **30.63**). Ясно, что в этом случае она не будет больше свободной от распределения (см. **30.36**), и в этом ее существенный недостаток по сравнению с X^2-критерием. Однако она имеет то преимущество, что позволяет строить доверительные интервалы для функции распределения при условии лишь, что последняя непрерывна.

*) При $\alpha \leqslant 0{,}2$ с большой точностью $d_\alpha \approx d_{\alpha/2}^+$, где d_α и d_α^+ — критические значения статистик D_n и D_n^+. Поэтому для получения процентных точек D_n^+ могут быть использованы таблицы процентных точек D_n (см. **30.55** и примечание на стр. 611). (*Прим. ред.*)

Вследствие сходимости $S_n(x)$ к истинной ф. р. $F(x)$ с вероятностью 1 (см. (30.98)) критерий D_n состоятелен против любой альтернативы $G(x) \neq F(x)$ *). Однако Мэсси (1950b, 1952) дал пример, в котором этот критерий смещен (см. упражнение 30.16). Он также установил нижнюю границу для мощности критерия при больших выборках, приводимую ниже.

30.60 Обозначим F_1 ф. р. при альтернативной гипотезе H_1. Пусть $F_0(x)$, как и ранее,— ф. р., подлежащая проверке, и

$$\Delta = \sup_x |F_1(x) - F_0(x)|. \tag{30.139}$$

Если d_α, как и ранее, — критическое значение D_n, то искомая мощность равна

$$P = \mathbf{P}\{\sup_x |S_n(x) - F_0(x)| > d_\alpha \mid H_1\}.$$

Это вероятность осуществления неравенства при каком-нибудь x. Ясно, что она не меньше, чем вероятность его осуществления при любом фиксированном значении x. Выберем то значение x_Δ, при котором F_0 и F_1 наиболее удалены, т. е.

$$\Delta = F_1(x_\Delta) - F_0(x_\Delta). \tag{30.140}$$

Таким образом, мы имеем

$$P \geqslant \mathbf{P}\{|S_n(x_\Delta) - F_0(x_\Delta)| > d_\alpha \mid H_1\},$$

или

$$P \geqslant 1 - \mathbf{P}\{F_0(x_\Delta) - d_\alpha \leqslant S_n(x_\Delta) \leqslant F_0(x_\Delta) + d_\alpha \mid H_1\}. \tag{30.141}$$

Далее, $S_n(x_\Delta)$ имеет биномиальное распределение с вероятностью $F_1(x_\Delta)$, равной вероятности попадания левее x_Δ. Таким образом, мы можем приблизить правую часть (30.141), пользуясь нормальным приближением к биномиальному распределению, т. е. асимптотически

$$P \geqslant 1 - (2\pi)^{-1/2} \int_{\frac{F_0 - F_1 - d_\alpha}{\{F_1(1-F_1)/n\}^{1/2}}}^{\frac{F_0 - F_1 + d_\alpha}{\{F_1(1-F_1)/n\}^{1/2}}} \exp\left(-\frac{1}{2}u^2\right) du, \tag{30.142}$$

причем в (30.142) и далее значения F_0 и F_1 берутся в точке x_Δ. Если F_1 задана, то (30.142) и есть искомая нижняя граница мощ-

*) Для утверждения о состоятельности критерия D_n соотношения (30.98) недостаточно, а требуется равномерная сходимость $S_n(x)$ к $F(x)$; с другой стороны, сходимость с вероятностью 1 можно заменить сходимостью по вероятности. Иначе говоря, требуется, чтобы $D_n \to 0$ по вероятности. (*Прим. ред.*)

ности. Ясно, что при $n \to \infty$ оба предела интегрирования возрастают. Если

$$d_\alpha < |F_0 - F_1| = \Delta, \tag{30.143}$$

то оба они стремятся к $+\infty$, когда $F_0 > F_1$, и к $-\infty$, когда $F_0 < F_1$. Таким образом, интеграл будет стремиться к нулю и мощность — к 1. Когда n возрастает, d_α стремится к нулю, так что (30.143) всегда рано или поздно будет выполнено. Следовательно, мощность $\to 1$, и критерий состоятелен.

Если F_1 не полностью задана, мы можем все же получить (худшую) нижнюю границу для мощности из (30.142). Так как $F_1(1-F_1) \leqslant 1/4$, мы имеем для достаточно больших n

$$P \geqslant 1 - (2\pi)^{-1/2} \int_{2n^{1/2}(F_0-F_1-d_\alpha)}^{2n^{1/2}(F_0-F_1+d_\alpha)} \exp\left(-\frac{1}{2}u^2\right) du,$$

и, пользуясь симметрией нормального распределения, при $F_0 < F_1$ это можно записать так:

$$P \leqslant 1 - (2\pi)^{-1/2} \int_{2n^{1/2}(\Delta-d_\alpha)}^{2n^{1/2}(\Delta+d_\alpha)} \exp\left(-\frac{1}{2}u^2\right) du. \tag{30.144}$$

Граница (30.144) зависит только от максимального отклонения Δ.

Бирнбаум (1953) получил близкие верхние и нижние границы для мощности D_n^+ в терминах Δ.

30.61 Используя (30.144) и вычисления, проделанные Уилльямсом (1950), Мэсси (1951) сравнивал значения Δ, при которых асимптотические мощности критериев X^2 и D_n достигают 0,5. При размере критерия $\alpha = 0{,}05$ критерий D_n может отличать с мощностью 0,5 отклонение Δ примерно вдвое меньшее, чем то, которое может отличать критерий X^2 с той же мощностью; даже при $n = 200$ отношение этих Δ равно 0,6, и оно все время убывает в пользу D_n, когда n растет. При $\alpha = 0{,}01$ критерии очень близки по своим качествам. Поскольку это сравнение основано на грубой нижней границе (30.144) для мощности D_n, мы должны заключить, что D_n — гораздо более чувствительный критерий для проверки согласия с непрерывным распределением.

Кац и др. (1955) указывают, что если используется процедура равных вероятностей Манна — Вальда из **30.28—29**, то критерию X^2 требуется Δ порядка $n^{-2/5}$, чтобы достичь мощности 1/2, в то время как для D_n требуется Δ порядка $n^{-1/2}$. Таким образом, требуемый объем выборки для D_n имеет асимптотически

порядок $n^{4/5}$ по сравнению с n для критерия X^2, и D_n асимптотически гораздо более эффективен — в действительности относительная эффективность X^2 будет стремиться к нулю при возрастании n.

Подробный обзор теории W^2, D_n и родственных им критериев дан Дарлингом (1957).

Вычисление D_n

30.62 Если мы строим доверительные пределы для неизвестной ф. р. $F(x)$, то не требуется никаких вычислений помимо простого расчета $S_n(x)$ и построения границ, отстоящих на $\pm d_\alpha$ от нее. Однако при использовании D_n для проверки гипотез нам придется находить $F_0(x)$ для каждого наблюденного значения x, что является довольно трудоемкой процедурой, даже когда $F_0(x)$ хорошо табулирована. Однако, поскольку статистикой критерия служит максимальное отклонение $S_n(x)$ от $F_0(x)$, часто бывает возможно посредством предварительного исследования данных выделить интервалы, в которых можно ожидать больших отклонений. Если начальные вычисления проделываются только для этих значений, то вычисления можно остановить, как только будет найдено одно отклонение, превосходящее d_α. (Такое сокращение вычислений невозможно для статистик, аналогичных W^2, которые зависят от всех отклонений.)

Дальнейшее заметное сокращение работы может быть достигнуто, как в следующем примере, принадлежащем Бирнбауму (1952).

Пример 30.6

Обрабатывается следующая выборка из 40 наблюдений, где наблюдения расположены в порядке возрастания: 0,0475; 0,2153; 0,2287; 0,2824; 0,3743; 0,3868; 0,4421; 0,5033; 0,5945; 0,6004; 0,6255; 0,6331; 0,6478; 0,7867; 0,8878; 0,8930; 0,9335; 0,9602; 1,0448; 1,0556; 1,0894; 1,0999; 1,1765; 1,2036; 1,2344; 1,2543; 1,2712; 1,3507; 1,3515; 1,3528; 1,3774; 1,4209; 1,4304; 1,5137; 1,5288; 1,5291; 1,5677; 1,7238; 1,7919; 1,8794.

Мы хотим проверить, с $\alpha = 0{,}05$, является ли ф. р. $F_0(x)$ этих наблюдений нормальной со средним 1 и дисперсией 1/6. По таблицам Бирнбаума (1952) для $n = 40$, $\alpha = 0{,}05$ находим $d_\alpha = 0{,}2101$. Рассмотрим наименьшее наблюдение $x_{(1)}$. Чтобы быть приемлемым, значение $F_0(x_{(1)})$ должно лежать между 0 и d_α, т. е. в интервале (0, 0,2101). Наблюденное значение $x_{(1)}$ равно 0,0475, и из таблиц нормальной ф. р. мы находим значение $F_0(x_{(1)}) = 0{,}0098$, входящее в указанный интервал, так что гипотеза не отвергается этим наблюдением. Далее, она не может

быть отвергнута последующими наблюдениями до тех пор, пока мы не дойдем до $x_{(i)}$, для которого либо (а) $i/40 - 0{,}2101 > 0{,}0098$, т. е. $i > 8{,}796$, либо (б) $F_0(x_{(i)}) > 0{,}2101 + 1/40$, т. е. $x_{(i)} > 0{,}7052$ (снова из таблиц). В правой части (б) прибавляется 1/40, потому что мы знаем, что $S_n(x_{(i)}) \geqslant 1/40$ при $i > 1$. Из наших данных имеем $x_{(i)} > 0{,}7052$ при $i > 14$. Следовательно, теперь нам нужно испытать $i = 9$ (из неравенства (а)). Здесь мы находим интервал принятия для $F_0(x_{(9)})$

$$(S_n(x_{(9)}) - d_\alpha,\ S_n(x_{(8)}) + d_\alpha) = (9/40 - 0{,}2101,\ 8/40 + 0{,}2101) = \\ = (0{,}0149,\ 0{,}4101).$$

По таблицам находим значение $F_0(x_{(9)}) = F_0(0{,}5945) = 0{,}1603$, являющееся приемлемым. Чтобы отвергнуть H_0, мы теперь требуем либо

$$i/40 - 0{,}2101 > 0{,}1603, \quad \text{т. е.} \quad i > 14{,}82,$$

либо

$$F_0(x_{(i)}) > 0{,}4101 + 1/40, \quad \text{т. е.} \quad x_{(i)} > 0{,}9052, \quad \text{т. е.} \quad i \geqslant 17.$$

Мы поэтому переходим к $i = 15$ и т. д. Читатель может проверить, что только шесть значений $i = 1, 9, 15, 21, 27, 34$ требуют в этом случае вычислений. Гипотеза принимается, так как в каждом из этих шести случаев значение F_0 лежит в доверительном интервале; она была бы отвергнута, а вычисления остановлены, если бы какое-нибудь значение оказалось вне этого интервала.

Критерии нормальности

30.63 В заключение этой главы мы кратко упомянем о задаче проверки нормальности, т. е. задаче проверки, принадлежит ли ф. р. выборки к семейству нормальных распределений, причем параметры не заданы заранее. Конечно, любой общий критерий согласия для сложной гипотезы может быть использован и для проверки нормальности, и в этом отношении никакого нового обсуждения не требуется. Однако распространенной практикой является проверка наблюденных моментных отношений b_1 и b_2 или простых функций от них на согласие с их распределениями, соответствующими гипотезе нормальности (см. **12.18** и упражнения 12.9, 12.10), и такие критерии иногда называют «критериями нормальности». Это очень свободное словоупотребление, и лучше называть их критериями асимметрии и эксцесса соответственно (см. **32.24** ниже). Гири (например, (1947)) разработал и исследовал иной критерий эксцесса, основанный на отношении выборочного среднего отклонения

к стандартному отклонению, который табулирован в *Biometrika Tables*, так же как и $\sqrt{b_1}$ и b_2 *).

Кац и др. (1955) рассматривают распределения D_n и W^2 для проверки нормальности, когда два параметра (μ, σ^2) оцениваются по выборке посредством $(\bar{x}, s^2)$. Предельные распределения свободны от параметров (поскольку это параметры сдвига и масштаба — см. **30.36**), но не получены в явном виде. Сообщается о некоторых выборочных экспериментах, которые дают эмпирические оценки для этих распределений.

Шапиро и Уилк (1965) построили новый критерий для проверки нормальности, основанный на регрессии порядковых статистик на их ожидаемые значения, пользуясь теорией из **19.18—20**, и провели обширные выборочные эксперименты для нахождения его распределения.

УПРАЖНЕНИЯ

30.1 Показать, что если при проверке сложной гипотезы используется несостоятельная совокупность оценок t, то статистика $X^2 \to \infty$ при $n \to \infty$.

(См. Фишер, 1924с.)

30.2 Пользуясь (30.33), показать, что матрица $\boldsymbol{M}$, определяемая в (30.37), при использовании вектора мультиномиальных МП-оценок $\hat{\boldsymbol{\theta}}$ сводится к

$$\boldsymbol{M} = \begin{pmatrix} \frac{1}{2}\boldsymbol{I} & \frac{1}{4}\boldsymbol{C} \\ \boldsymbol{C}^{-1} & \frac{1}{2}\boldsymbol{I} \end{pmatrix}$$

и что $\boldsymbol{M}$ идемпотентна с $\operatorname{tr} \boldsymbol{M} = s$. Этим подтверждается результат **30.10** и **30.14**, что X^2 распределен асимптотически, как χ^2_{k-s-1}, когда используется $\hat{\boldsymbol{\theta}}$.

(Ватсон, 1959.)

30.3 Показать, исходя из предельной совместной нормальности n_i, что при $n \to \infty$ дисперсия статистики X^2 для простой гипотезы в случае равных вероятностей $(p_{0i} = 1/k)$ равна

$$\mathrm{D}(X^2) = 2k^2 \left\{ \sum_i p_{1i}^2 - \left(\sum_i p_{1i}^2 \right)^2 \right\} + 4(n-1)\, k^2 \left\{ \sum_i p_{1i}^3 - \left(\sum_i p_{1i}^2 \right)^2 \right\},$$

где p_{1i}, $i = 1, 2, \ldots, k$, — истинные вероятности классов. Проверить, что это сводится к значению $2(k-1)$, когда

$$p_{1i} = p_{0i} = 1/k.$$

(Манн и Вальд, 1942.)

30.4 Получить результат из **30.27** о нецентральном χ^2 для распределения статистики X^2 при альтернативной гипотезе.

(См. Кокрэн, 1952.)

---

*) Эти таблицы воспроизведены в *Таблицах математической статистики*. (*Прим. ред.*)

30.5 Показать, исходя из моментов мультиномиального распределения (см. (5.80)), что точная дисперсия статистики X^2 для простой гипотезы дается формулой

$$n\mathbf{D}(X^2) = 2(n-1)\left\{2(n-2)\sum_i \frac{p_{1i}^3}{p_{0i}^2} - (2n-3)\left(\sum_i \frac{p_{1i}^2}{p_{0i}}\right)^2 - \right.$$

$$\left. - 2\left(\sum_i \frac{p_{1i}^2}{p_{0i}}\right)\left(\sum_i \frac{p_{1i}}{p_{0i}}\right) + 3\sum_i \frac{p_{1i}^2}{p_{0i}^2}\right\} - \left(\sum_i \frac{p_{1i}}{p_{0i}}\right)^2 + \sum_i \frac{p_{1i}}{p_{0i}^2}.$$

(Патнайк, 1949.)

30.6 Для той же альтернативной гипотезы, что и в примере 30.3, а именно гамма-распределения с параметром 1,5, с помощью *Biometrika Tables* получить значения p_{1i} для группировки на четыре класса с неравными вероятностями примера 30.2. Подсчитать параметр нецентральности (30.62) для этого случая и показать сравнением с примером 30.3, что группировка с неравными вероятностями потребовала бы примерно на 25% большего объема выборки, чем группировка с равными вероятностями, чтобы достичь той же мощности при этой альтернативе.

30.7 k независимых стандартных нормальных величин x_j подчинены c однородным линейным ограничениям. Показать, что $S = \sum_{j=1}^{k} x_j^2$ распределено независимо от знаков x_j. Показать, что если $c = 1$ и это ограничение есть $\sum_{j=1}^{k} x_j = 0$, то все последовательности знаков равновероятны (за исключением двух: все положительны и все отрицательны, которые не могут произойти), но что это, вообще говоря, не так при $c > 1$. Вывести отсюда, что любой критерий, основанный на последовательности знаков отклонений $(n_i - np_{0i})$ наблюденных частот от гипотетических, асимптотически независим от критерия X^2, когда справедлива гипотеза H_0.

(Ф. Дэвид, 1947; Сил, 1948; Фрэзер, 1950.)

30.8 M элементов одного вида и N элементов другого расположены в случайном порядке $(M, N > 0)$. *Серия* определяется как подпоследовательность элементов одного вида, непосредственно перед которой и после которой стоят элементы другого вида. Пусть R — число серий во всей последовательности $(2 \leqslant R \leqslant M + N)$. Показать, что

$$\mathbf{P}\{R = 2s\} = 2\binom{M-1}{s-1}\binom{N-1}{s-1}\Big/\binom{M+N}{M},$$

$$\mathbf{P}\{R = 2s-1\} = \left\{\binom{M-1}{s-2}\binom{N-1}{s-1} + \binom{M-1}{s-1}\binom{N-1}{s-2}\right\}\Big/\binom{M+N}{M}$$

и что

$$\mathbf{M}(R) = 1 + \frac{2MN}{M+N}, \qquad \mathbf{D}(R) = \frac{2MN(2MN-M-N)}{(M+N)^2(M+N-1)}.$$

(Стивенс (1939); Вальд и Волфовиц (1940). Свед и Эйзенхарт (1943) табулировали распределение R для $M \leqslant N \leqslant 20$.)

30.9 Пользуясь упражнениями 30.7 и 30.8, показать, что если имеются M положительных и N отрицательных отклонений $(n_i - np_{0i})$, то мы можем использовать критерий серий в дополнение к критерию X^2 для проверки простой гипотезы. Пользуясь упражнением 16.4, показать, что если P_1 — вероятность получить значение X^2 не меньшее, чем наблюденное, а P_2 — вероятность получить значение R не большее, чем наблюденное, то $U = -2(\log P_1 + \log P_2)$ распределено асимптотически, как χ^2 с 4 степенями свободы, и большие значения U образуют критическую область составного критерия.

(Ф. Дэвид, 1947.)

30.10 Пусть $x_1, x_2, \ldots, x_n$ — независимые случайные величины с одним и тем же распределением $f(X|\theta_1, \theta_2)$. θ_1 и θ_2 оцениваются статистиками $t_1(x_1, x_2, \ldots, x_n)$, $t_2(x_1, x_2, \ldots, x_n)$. Показать, что случайные величины

$$y_i = \int_{-\infty}^{x_i} f(u \mid t_1, t_2)\, du$$

не независимы и их распределение зависит, вообще говоря, от f, θ_1 и θ_2; но что если θ_1 и θ_2 — параметры сдвига и масштаба соответственно, а t_1, t_2 — в подходящем смысле инвариантные оценки, то распределение y_i не зависит от θ_1 и θ_2, а только от вида f.

(Ф. Дэвид и Джонсон, 1948.)

30.11 Показать, что при проверке сложной гипотезы статистика X^2, использующая мультиномиальные МП-оценки, асимптотически эквивалентна статистике критерия отношения правдоподобия, когда справедлива H_0.

30.12 Показать, что статистика критерия согласия Неймана (30.80) эквивалентна при больших выборках критерию отношения правдоподобия для простой гипотезы (30.78).

(Пирсон, 1938.)

30.13 Проверить значения среднего и дисперсии (30.81), (30.82).

30.14 Доказать формулу (30.102) для дисперсии W^2.

30.15 Проверить, что nW^2 может быть представлено в виде (30.105).

30.16 Показать с помощью чертежа, что в задаче проверки простой гипотезы, задающей ф. р. $F_0(x)$, при простой альтернативе $F_1(x)$, удовлетворяющей условиям

$$F_1(x) < F_0(x), \quad \text{когда} \quad F_0(x) < d_\alpha,$$
$$F_1(x) = F_0(x) \quad \text{в остальных случаях},$$

критерий D_n (с критическим значением d_α) может быть смещенным.

(Мэсси, 1950b, 1952.)

30.17 Берутся n наблюдений u_r, имеющих равномерное распределение на интервале $(0, 1)$, которые разбивают этот интервал на $n+1$ отрезков c_j, где $c_j \geqslant 0$ и $\sum_{j=1}^{n+1} c_j = 1$. Эти c_j упорядочиваются так, что $c_{(1)} \leqslant c_{(2)} \leqslant \ldots \ldots \geqslant c_{(n+1)}$. Показать, что неотрицательные случайные величины

$$g_1 = (n+1)\, c_{(1)},$$

$$g_j = (n+2-j)\,(c_{(j)} - c_{(j-1)}), \quad j = 2, 3, \ldots, n+1; \quad \sum_{j=1}^{n+1} g_j = 1,$$

имеют распределение $dF_g = n!\, dg_1 \ldots dg_n$ и что неупорядоченные c_j имеют то же самое распределение, так что $dF_c = n!\, dc_1 \ldots dc_n$ (при этом $(n+1)$-я величина в обоих случаях опущена, чтобы избежать сингулярности распределения). Вывести отсюда, что величины

$$w_r = \sum_{j=1}^{r} g_j, \quad r = 1, 2, \ldots, n,$$

распределены в точности как порядковые статистики $u_{(r)}$, $r = 1, 2, \ldots, n$ первоначальной выборки. Таким образом, любой критерий согласия, основанный на вероятностном интегральном преобразовании, может быть применен к w_r, так же как и к $u_{(r)}$, полученным в результате преобразования.

(См. Дёрбин (1961), где было найдено с помощью выборочных экспериментов, что односторонний критерий Колмогорова D_n^-, примененный к w_r, имеет лучшие свойства мощности, чем обычный двусторонний критерий D_n, при выявлении изменения формы распределения.)

30.18 Пусть θ — неизвестные параметры в задаче проверки сложной гипотезы согласия для n наблюдений $\boldsymbol{x}$. Предположим, что $\boldsymbol{t}$ — минимальная достаточная статистика для θ, имеющая менее чем n компонент, и что имеется взаимно однозначное преобразование $\boldsymbol{x}$ в $(\boldsymbol{t}, \boldsymbol{u})$, где $\boldsymbol{u}$ не зависит от $\boldsymbol{t}$. Показать, что если значение $\boldsymbol{t}$ отбросить и заменить случайным наблюдением $\boldsymbol{t}'$ с распределением, отвечающим известному значению θ, то совокупность наблюдений $\boldsymbol{x}'$, полученная обратным преобразованием из $(\boldsymbol{t}', \boldsymbol{u})$, имеет распределение, не зависящее от θ, так что гипотеза согласия становится простой.

(Дёрбин, 1961.)

ГЛАВА 31

УСТОЙЧИВЫЕ И СВОБОДНЫЕ ОТ РАСПРЕДЕЛЕНИЯ ПРОЦЕДУРЫ

31.1 В ходе рассмотрения различных аспектов статистической теории мы во многих случаях обнаруживали, что наибольший прогресс может быть достигнут тогда, когда наблюдения получены из нормальной генеральной совокупности. Основная причина этого — сферическая симметрия, характеризующая нормальность, но нас сейчас интересует не это. Вопрос, который мы собираемся обсуждать, состоит в том, насколько оправданным может быть применение этой так называемой «нормальной теории» в ситуации, когда распределения наблюдений на самом деле не нормальны. Потому что в свете относительного обилия теоретических результатов, относящихся к нормальному случаю, возникает искушение рассматривать распределения как нормальные, если только не доказано противное, и использовать стандартную нормальную теорию везде, где это возможно. Возникает вопрос, насколько вероятно, что такие оптимистические предположения нормальности могут привести к серьезным ошибкам.

Для задач проверки гипотез мы можем сформулировать проблему более точно в стиле нашего обсуждения подобных областей в **23.4**. Напомним, что там нашей целью было установить размер критерия равным α независимо от значений некоторых мешающих параметров. Теперь нас интересует вопрос, по существу, такого же рода, но относящийся к функциональной форме распределения, а не к его неизвестным параметрам: чувствителен ли размер критерия α к изменениям формы распределения?

Статистическая процедура, нечувствительная к отклонениям от предположений, лежащих в ее основе, называется «устойчивой» *). Исследования устойчивости проводились многими авторами. Большое количество работ относится к дисперсионному анализу, и мы отложим обсуждение этого до 3-го тома. Сейчас мы ограничимся результатами, связанными с процедурами, которые нам уже встречались. Бокс и Андерсен (1955) дают общий обзор этого предмета.

*) В оригинале — «robust». Фраза в оригинале продолжается: «удачный термин, введенный Боксом (1953) и теперь имеющий всеобщее употребление». В русской литературе установившегося термина нет. (*Прим. ред.*)

Устойчивость стандартных процедур «нормальной теории»

31.2 Начавшись с ранних экспериментальных исследований, в первую очередь Э. Пирсона, изучение устойчивости продолжалось теоретическими исследованиями, среди которых работы Бартлетта (1935a), Гири (1936, 1947) и Гейена (1949—1951), по существу, имеют аналогичную форму. Предполагается, что наблюдения получены из генеральных совокупностей, определяемых разложениями Грама — Шарлье или Эджворта, и поправочные члены к нормальной теории получаются в виде функций от нормированных старших семиинвариантов, в частности $\varkappa_3$ и $\varkappa_4$. Эти результаты могут быть, в общем, выражены утверждением, что критерии, касающиеся генеральных средних (т. е. t-критерии Стьюдента для среднего значения нормального распределения и для разности средних двух нормальных распределений с одинаковыми дисперсиями), довольно нечувствительны к отклонениям от нормальности, а критерии, касающиеся дисперсий (т. е. χ^2-критерий для дисперсии нормального распределения, F-критерий для отношения дисперсий двух нормальных распределений и модифицированный критерий отношения правдоподобия для равенства нескольких дисперсий в примерах 24.4, 24.6), очень чувствительны к таким отклонениям. Критерии для средних устойчивы; в сравнении с ними критерии для дисперсий следует считать неустойчивыми. У нас нет здесь места для подробного вывода этих результатов, но их легко объяснить в общих словах.

31.3 Ключевым моментом при выводе распределения t-статистики Стьюдента является независимость ее числителя и знаменателя, которая выполняется точно только для нормальных исходных распределений. Если мы производим выборку из генеральной совокупности, отличной от нормальной, то центральная предельная теорема тем не менее дает нам, что выборочное среднее и несмещенная оценка дисперсии $s^2 = k_2$ будут распределены асимптотически нормально. Более того, мы знаем из правила 10 для выборочных семиинвариантов k-статистик в **12.14**, что

$$\varkappa(21) = \varkappa_3/n, \tag{31.1}$$

$$\varkappa(2^r1^s) = O\left(n^{-(r+s-1)}\right). \tag{31.2}$$

Так как

$$\varkappa(1^2) = \frac{\varkappa_2}{n}, \quad \varkappa(2^2) = \frac{\varkappa_4}{n} + \frac{2\varkappa_2^2}{n-1},$$

то для асимптотической корреляции между $\bar{x}$ и s^2 мы имеем из (31.1)

$$\rho = \varkappa_3/\{\varkappa_2(\varkappa_4 + 2\varkappa_2^2)\}^{1/2}. \tag{31.3}$$

Если генеральная совокупность симметрична, то $\varkappa_3$ и ρ в (31.3) равны нулю, и, следовательно, $\bar{x}$ и s^2 асимптотически независимы, так что нормальная теория будет справедлива для достаточно больших n. Если $\varkappa_3 \neq 0$, то (31.3) будет меньше, когда $\varkappa_4$ велико, но останется не равным нулю. Ситуацию, однако, спасает тот факт, что точное t-распределение Стьюдента само приближается к нормальному при $n \to \infty$, так же как и, по центральной предельной теореме, распределение статистики

$$t = (\bar{x} - \mu)/(s^2/n)^{1/2}, \tag{31.4}$$

поскольку s^2 сходится по вероятности к σ^2. Оба предельных распределения совпадают.

Таким образом, каково бы ни было исходное распределение, статистика (31.4) асимптотически нормальна, и, следовательно, в пределе справедлива нормальная теория. Если исходное распределение симметрично, мы можем ожидать более быстрой сходимости к распределению, даваемому нормальной теорией (t-распределение Стьюдента). Это на самом деле и есть результат, подтверждаемый подробными исследованиями: для малых выборок нормальная теория менее устойчива по отношению к отклонениям асимметрии, чем к отклонениям эксцесса.

31.4 Аналогично обстоит дело для двухвыборочной t-статистики Стьюдента. Если две выборки получены из одной и той же генеральной совокупности, отличной от нормальной, а мы применяем нормальную статистику

$$t = \{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)\} \Bigg/ \left\{ \left[\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \right] \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \right\}^{1/2}, \tag{31.5}$$

то ковариация между $(\bar{x}_1 - \bar{x}_2)$ и членом в квадратных скобках в знаменателе, который обозначим s^2, дается формулой

$$\mathrm{cov} = \varkappa_3 \left\{ \frac{n_1 - 1}{n_1 + n_2 - 2} \cdot \frac{1}{n_1} - \frac{n_2 - 1}{n_1 + n_2 - 2} \cdot \frac{1}{n_2} \right\} = \frac{\varkappa_3}{n_1 + n_2 - 2} \left(\frac{1}{n_2} - \frac{1}{n_1} \right),$$

тогда как соответствующие дисперсии равны

$$\mathrm{D}(\bar{x}_1 - \bar{x}_2) = \varkappa_2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right), \qquad \mathrm{D}(s^2) \sim (\varkappa_4 + 2\varkappa_2^2)/(n_1 + n_2).$$

Корреляция, следовательно, равна асимптотически

$$\rho = \frac{\varkappa_3}{\{\varkappa_2(\varkappa_4 + 2\varkappa_2^2)\}^{1/2}} \frac{(n_1 n_2)^{1/2}}{n_1 + n_2 - 2} \left(\frac{1}{n_2} - \frac{1}{n_1} \right). \tag{31.6}$$

Снова, если $\varkappa_3=0$, асимптотическая нормальность влечет асимптотическую независимость. Мы также видим, что ρ равно нулю, если $n_1 = n_2$. Во всяком случае, когда n_1 и n_2 становятся большими, центральная предельная теорема приводит к асимптотической нормальности выражения (31.5) и, следовательно, к согласию с t-распределением Стьюдента.

И опять это в точности совпадает с результатами, полученными Бартлеттом (1935а) и Гейеном (1949—1951): если объемы выборок равны, то даже асимметрия исходного распределения вызывает малое отклонение от нормальной теории. Если исходное распределение симметрично, то критерий будет устойчив даже при неодинаковых объемах выборок.

31.5 Для t-критериев Стьюдента изучались также влияния более сложных отклонений от нормальности. Хюрениус (1950) рассматривал выборку из составного нормального распределения; другие шведские авторы, наиболее поздним из которых является Закриссон (1959), дающий ссылки на более ранние работы, также рассматривали различные формы генеральных совокупностей, составленных из нормальных подсовокупностей. Роббинс (1948) нашел распределение t, когда наблюдения получены из нормальных распределений, отличающихся только средними. Для двухвыборочного критерия Гири (1947) и Гейен (1949—1951) рассматривают случай, когда две выборки могут происходить из различных генеральных совокупностей.

31.6 Когда мы переходим к критериям для дисперсий, картина становится совсем иной. Ключевым пунктом нормальной теории во всех критериях для дисперсий является тот факт, что отношение $z = \sum_{i=1}^{n} (x_i - \bar{x})^2/\sigma^2$ распределено как χ^2 с $n-1$ степенями свободы. Если мы рассмотрим выборочные семиинварианты величины $k_2 = \varkappa_2 z/(n-1)$, то из (12.35) мы видим, что

$$\mathsf{D}(z) = \left(\frac{n-1}{\varkappa_2}\right)^2 \varkappa(2^2) = 2(n-1) + \frac{(n-1)^2}{n}\frac{\varkappa_4}{\varkappa_2^2} \sim (n-1)\left(2 + \frac{\varkappa_4}{\varkappa_2^2}\right), \tag{31.7}$$

тогда как из (12.36)

$$\mu_3(z) = \left(\frac{n-1}{\varkappa_2}\right)^3 \varkappa(2^3) =$$

$$= \left(\frac{n-1}{\varkappa_2}\right)^3 \left\{\frac{\varkappa_6}{n^2} + \frac{12\varkappa_4\varkappa_2}{n(n-1)} + \frac{4(n-2)\varkappa_3^2}{n(n-1)^2} + \frac{8\varkappa_2^3}{(n-1)^2}\right\} \sim$$

$$\sim (n-1)\left\{\frac{\varkappa_6}{\varkappa_2^3} + \frac{12\varkappa_4}{\varkappa_2^2} + \frac{4\varkappa_3^2}{\varkappa_2^3} + 8\right\},$$

и аналогично для старших моментов из (12.37)—(12.39). Из этих выражений становится ясно, что распределение z зависит

от всех нормированных семиинвариантных отношений $\varkappa_3^2/\varkappa_2^3$, $\varkappa_4/\varkappa_2^2$ и т. д. и что члены, содержащие эти отношения, имеют тот же порядок по n, что и постоянные члены нормальной теории. Эти дополнительные члены пропадают тогда и только тогда, когда все высшие семиинварианты равны нулю, т. е. когда распределение наблюдений нормально. В противном случае (31.7) показывает, что, хотя z распределено асимптотически нормально, распределение z не будет асимптотически приближаться к χ^2-распределению, даваемому нормальной теорией. Здесь центральная предельная теорема не спасает нас, так как распределение z стремится к нормальному распределению, отличному от того, которого мы хотим.

31.7 Поскольку $\varkappa_4$ фигурирует в (31.7), а $\varkappa_3$ — нет, следует ожидать, что отклонение от нормального эксцесса будет оказывать большее влияние на распределение, и именно в этом состоит результат, полученный после детальных выкладок Гейеном (1949—1951) для критериев χ^2 и отношения дисперсий. Бокс (1953) нашел, что расхождения с нормальной теорией увеличиваются, когда число сравниваемых дисперсий возрастает, и его рассуждения достаточно просты, чтобы воспроизвести их здесь.

Предположим, что k выборок объемов n_i $(i = 1, 2, \ldots, k)$ извлечены из генеральных совокупностей, имеющих одну и ту же дисперсию $\varkappa_2$ и один и тот же коэффициент эксцесса $\gamma_2 = \varkappa_4/\varkappa_2^2$.

Из (31.7) мы тогда имеем асимптотически для любой из выборок

$$\mathrm{D}(s_i^2) = 2\varkappa_2^2\left(1 + \tfrac{1}{2}\gamma_2\right)/n_i, \tag{31.8}$$

где s_i^2 — несмещенная оценка для $\varkappa_2$. По центральной предельной теореме s_i^2 асимптотически нормальна со средним $\varkappa_2$ и дисперсией (31.8) и, следовательно, распределена так, как если бы она была получена из нормальной генеральной совокупности и основана на $N_i = n_i/\left(1 + \frac{1}{2}\gamma_2\right)$ наблюдениях вместо n_i. Таким образом, воздействие на модифицированный критерий ОП для сравнения k нормальных дисперсий, приведенный в (24.44), состоит в том, что $-2\log l^*/\left(1 + \frac{1}{2}\gamma_2\right)$, а не само $-2\log l^*$ распределено асимптотически как χ^2 с $k - 1$ степенями свободы.

Влияние этой поправки на распределение, даваемое нормальной теорией, может быть чрезвычайно сильным. В следующей таблице мы приводим некоторые результаты вычислений Бокса (1953).

Истинная вероятность превышения критического уровня, даваемого асимптотической нормальной теорией для $\alpha = 0{,}05$

γ_2 \ k	2	3	5	10	30
−1	0,0056	0,0025	0,0008	0,0001	$0{,}0^61$
0	0,05	0,05	0,05	0,05	0,05
1	0,110	0,136	0,176	0,257	0,498
2	0,166	0,224	0,315	0,489	0,849

Как видно из таблицы, отклонение от значения 0,05, соответствующего нормальной теории, возрастает с ростом $|\gamma_2|$ и с ростом k при любом фиксированном $\gamma_2 \neq 0$.

31.8 Хотя результат **31.7** имеет асимптотический характер, Бокс (1953) показал, что аналогичные расхождения происходят и при малых выборках. Поразительное отсутствие устойчивости дисперсионного критерия заставило его рассматривать статистику l^* из (24.44) как статистику критерия для эксцесса, и он нашел, что чувствительность этого критерия имеет тот же порядок, что и у общеупотребительных критериев, указанных в **30.63.**

31.9 Наконец, мы кратко упомянем, что Гейен (1949—1951) рассмотрел устойчивость выборочного коэффициента корреляции r и фишеровского z-преобразования r к отклонениям от двумерной нормальности. Когда теоретический коэффициент корреляции $\rho = 0$ и, в частности, когда случайные величины независимы, распределение r устойчиво даже при столь малом объеме выборки, как 11; но при больших значениях ρ отклонения от нормальной теории становятся заметными. Когда исходное распределение отлично от нормального, z-преобразование остается асимптотически нормальным, но скорость сходимости уменьшается. На среднее и дисперсию z, с точностью до порядка n^{-1}, не влияет асимметрия исходных маргинальных распределений, но влияние отклонений от нормального эксцесса может быть заметным; в частности, дисперсия z чувствительна к форме исходного распределения даже при больших выборках, хотя математическое ожидание z медленно приближается к своему нормальному значению, когда n растет.

Хотеллинг (1961) применяет совершенно иной подход к задачам устойчивости и приводит полезный библиографический список.

Хубер (1964) проводит общее исследование устойчивости оценок параметра сдвига; см. также Бикел (1965) и Гастверт (1966).

Нормализующие преобразования

31.10 Исследование устойчивости имеет своей целью выяснение области применимости стандартных процедур нормальной теории. Как мы видели, эта область может быть широкой или чрезвычайно узкой, но на практике часто трудно решить, будут ли стандартные процедуры приближенно справедливы или приведут к неверному результату. Мы сейчас обсудим два других

подхода к ситуациям, когда не выполнены предположения нормальности.

Первая возможность состоит в том, чтобы искать преобразование, которое сделало бы распределение наблюдений более близким к нормальному, так что к преобразованным наблюдениям была бы применима нормальная теория. Это преобразование может иметь вид, рассмотренный в **6.25—26**, где мы нормализовали подбором полиномиального преобразования. Либо мы сможем найти простое нормализующее функциональное преобразование (см. **6.27—35** и фишеровское z-преобразование коэффициента корреляции в (16.75)). Трудность в обоих случаях состоит в том, что мы должны заранее знать первоначальное распределение, чтобы найти, какое преобразование лучше применить; такая информация может оказаться доступной в теоретическом контексте, таком, как исследование выборочного распределения статистики, но ее гораздо труднее получить, когда интересующее нас распределение возникает в экспериментальной работе.

По счастью, преобразования, предназначенные для того, чтобы стабилизировать дисперсию (т. е. чтобы сделать ее не зависящей от некоторого параметра), часто служат и для нормализации распределения, к которому они применяются, — примером этого может служить фишеровское z-преобразование коэффициента r. Упражнение 16.18 показывает, как знание связи между средним и дисперсией распределения позволяет произвести простое преобразование, стабилизирующее дисперсию. Такие преобразования чаще всего используются в дисперсионном анализе, и мы отложим их подробное обсуждение до изложения этой темы в 3-м томе.

Свободные от распределения процедуры

31.11 Второй из альтернативных подходов, упомянутых в начале **31.10**, будет радикальным. Вместо того, чтобы держаться за стандартные методы нормальной теории (либо потому, что они устойчивы и приближенно применимы в случае отклонения от нормальности, либо потому, что они становятся приближенно применимы после преобразования наблюдений), мы теперь полностью отказываемся от них и ищем новый подход к нашим проблемам. Можем ли мы найти статистические процедуры, которые были бы применимы в широком классе исходных распределений, скажем в классе всех непрерывных распределений? Если можем, то они с необходимостью будут справедливы и для нормальных распределений и наша устойчивость будет точной и гарантированной. Такие процедуры называются свободными от распределения, как мы уже видели в **30.35**, поскольку их

применимость вовсе не зависит от формы исходных распределений, если только они непрерывны.

Остаток этой главы и некоторые разделы двух последующих глав будут посвящены свободным от распределения методам. Прежде всего мы обсудим, в каком отношении находятся свободные от распределения методы к схеме «параметрический — непараметрический», введенной в **22.3.**

31.12 Ясно, что, если мы имеем дело с параметрической задачей (например, проверкой параметрической гипотезы или оценкой параметра), метод, которым мы пользуемся, может быть и может не быть свободным от распределения. По-видимому, не столь же очевидно, что даже в непараметрической задаче метод также может быть и может не быть свободным от распределения. Например, в главе 30 мы рассматривали критерии для сложной гипотезы согласия, представляющей собой непараметрическую задачу, и нашли, что статистика критерия не является даже асимптотически свободной от распределения в общем случае, когда применяются не мультиномиальные МП-оценки. Опять же, если мы используем выборочное моментное отношение $b_2 = m_4/m_2^2$ как критерий нормальности, то это — непараметрическая задача, но распределение b_2 сильно зависит от формы исходного распределения.

Однако большинство свободных от распределения процедур было предназначено для непараметрических задач, таких, как проверка гипотезы, что два непрерывных распределения равны между собой, и поэтому в литературе установилась довольно произвольная взаимозаменяемость терминов «непараметрический» и «свободный от распределения». Мы будем всегда употреблять их в тех совершенно различных значениях, которые мы определили: «непараметрический» относится к задаче, а «свободный от распределения» — к методу, используемому для решения задачи.

Свободные от распределения методы для непараметрических задач

31.13 Ниже следуют основные классы непараметрических задач, которые могут быть решены свободными от распределения методами.

(1) *Задача двух выборок.* Проверяемая гипотеза состоит в том, что две генеральные совокупности, из которых мы имеем случайные выборки наблюдений, одинаковы.

(2) *Задача k выборок.* Это обобщение (1) на $k > 2$ генеральных совокупностей.

(3) *Случайность.* Имеется ряд из n наблюдений над одной случайной величиной, упорядоченной каким-то образом (обычно

по времени). Проверяется гипотеза, что наблюдения получены независимо между собой из одной и той же генеральной совокупности.

(4) *Независимость в двумерной генеральной совокупности.* Проверяется гипотеза, что двумерное распределение распадается на два независимых маргинальных распределения.

Все это — задачи проверки гипотез, и, действительно, дело обстоит так, что большинство свободных от распределения методов относится к проверке гипотез, а не к оцениванию. Однако мы можем находить свободные от распределения

(1а) *Доверительные интервалы для различия в сдвиге двух непрерывных распределений, в остальном одинаковых,*

(5) *Доверительные интервалы и критерии для квантилей,*

и

(6) *Толерантные интервалы для непрерывного распределения.*

В главе 30 мы уже рассматривали

(7) *Свободные от распределения критерии согласия,*

и

(8) *Доверительные интервалы для непрерывной функции распределения.*

Перечисленные выше категории содержат основную массу работы, проделанной до сих пор по не зависящим от распределения методам, хотя, как мы увидим, они и не исчерпывающи.

Очень полная библиография по этому предмету дана Сэвиджем (1962).

31.14 Читатель, вероятно, заметил, что задачи (1)—(3) в **31.13** однотипны, поскольку они касаются проверки тождественности нескольких одномерных непрерывных распределений, и у него, возможно, возник вопрос, почему с ними объединена задача (4). Причина этого в том, что, видоизменяя задачу (4), можно прийти к задачам (1)—(3). Мы здесь кратко опишем эту связь, оставляя подробности до тех пор, пока не перейдем к конкретным критериям.

Предположим, что в задаче (3) мы приписываем значениям x числа, характеризующие порядок их расположения, и рассматриваем эти числа как наблюдения над случайной величиной y. Задача (3) тогда сводится к проверке независимости x и y, т. е. к частному случаю задачи (4). Предположим теперь, что в задаче (4) мы разбили на две части область значений второй случайной величины, скажем z, и что мы полагаем $y = 1$ или 2 в зависимости от того, в какую из частей попадает наблюдение z. Если мы теперь будем проверять независимость x и y, то задача (4) сведется к задаче (1), так как если x не зависит от y-классификации, то распределения x для $y = 1$ и

для $y = 2$ должны быть тождественны. Аналогично мы сводим задачу (4) к задаче (2), разбивая область значений z на $k > 2$ классов, придавая значения $y = 1, 2, \ldots, k$ и проверяя независимость x и y.

Построение свободных от распределения критериев

31.15 Как строить свободные от распределения критерии для непараметрических задач? При обсуждении критериев согласия в главе 30 мы уже встречали два метода: один состоял в применении вероятностного интегрального преобразования, которое для простых гипотез дает свободный от распределения критерий; второй состоял в сведении задачи к задаче о мультиномиальном распределении, как в случае критерия X^2, — в следующей главе мы увидим, что этот последний способ в своей простейшей форме служит для построения критерия (так называемого критерия знаков) для задачи (5) из **31.13**. Но важнейшие классы свободных от распределения критериев для задач (1)—(4) основываются на других методах, которые мы теперь рассмотрим.

Если мы ничего не знаем о форме исходных распределений, кроме, возможно, того, что они непрерывны, мы, очевидно, не можем найти подобных областей в пространстве выборок теми методами, которые применялись для параметрических задач в главе 23. Однако продвижение все же возможно. Прежде всего мы произведем некоторые необходимые изменения в наших определениях достаточности и полноты.

При отсутствии параметрической формулировки мы должны связывать эти определения непосредственно с исходными ф. р.; в то время как раньше мы называли статистику t достаточной для параметра θ, если была возможна факторизация (17.68), мы теперь определим семейство C распределений, и пусть θ — просто некоторая переменная, которая служит индексом для членов этого семейства. Имея это в виду, будем называть статистику t достаточной для семейства C, если для всех θ имеет место факторизация (17.68). Аналогично, определения полноты и ограниченной полноты семейства распределений в **23.9** остаются в силе для непараметрических ситуаций, если в качестве θ взять индексную переменную для членов семейства.

31.16 В примерах 23.5 и 23.6 мы видели, что совокупность порядковых статистик $t = (x_{(1)}, x_{(2)}, \ldots, x_{(n)})$ является достаточной статистикой в некоторых параметрических задачах, хотя и не обязательно минимальной достаточной статистикой. Интуитивно очевидно, что t всегда будет достаточной статистикой, если все наблюдения получены из одного и того же распределения, так как в этом случае мы не теряем никакой информа-

ции, упорядочивая наблюдения. (Очевидно также, что t будет минимальной достаточной статистикой, если о форме исходного распределения ничего не известно.) Если исходное распределение непрерывно, мы отмечали в **23.5**, что подобные области могут всегда быть построены посредством перестановки координат пространства выборок для критериев, размер которых кратен $(n!)^{-1}$. Такая перестановка оставляет совокупность порядковых статистик без изменений. Если относительно формы исходного распределения не известно ничего, то ясно, что мы не можем получить подобных областей никаким другим способом. Поэтому из результата **23.19** следует, что совокупность порядковых статистик ограниченно полна для семейства всех непрерывных ф. р. *).

Следовательно, мы видим, что, если нам нужно строить подобные критерии для таких гипотез, как в задачах (1)—(4) в **31.13**, мы должны пользоваться *критериями перестановок*, которые основываются, по существу, на том факте, доказанном в **11.4** и очевидном из симметрии, что каждое упорядочение выборки из непрерывной ф. р. имеет одну и ту же вероятность $(n!)^{-1}$. При этом еще остается вопрос, каким критерием перестановок следует пользоваться для данной конкретной гипотезы.

Эффективность свободных от распределения критериев

31.17 Поиск свободных от распределения процедур вызван желанием расширить область применимости наших выводов. Мы не можем ожидать, что большой выигрыш в общности будет достигнут без некоторой потери эффективности в конкретной ситуации; иными словами, мы не можем ожидать, что свободный от распределения критерий, выбранный без знания формы исходного распределения, будет так же эффективен, как и критерий, которым мы пользовались бы, если бы знали это исходное распределение. Но использовать это как аргумент против свободных от распределения процедур было бы явно ошибочным: именно отсутствие информации относительно исходного распределения заставляет нас пользоваться свободными от распределения методами. Единственная «справедливая» мера эффективности для свободного от распределения критерия — это та, которая дается другими свободными от распределения критериями. Мы должны, естественно, выбирать наиболее эффективный из имеющихся таких критериев.

Но в каком смысле должны мы говорить об эффективности? Даже в параметрическом случае РНМ критерии редки, и мы

*) То, что она на самом деле полна, доказано непосредственно, например, в книге Лемана (1959); этот результат принадлежит Шеффе (1943b).

не можем надеяться найти свободные от распределения критерии, которые были бы наиболее мощными против всех возможных альтернатив. Мы, таким образом, приходим к тому, чтобы исследовать мощность свободных от распределения критериев против параметрических альтернатив к проверяемым непараметрическим гипотезам. Несмотря на парадоксальное звучание, в этом нет никакого противоречия, и такая процедура имеет одно большое практическое достоинство. Если мы исследуем мощность против альтернатив, рассматриваемых в теории нормального распределения, мы получаем меру того, как много мы можем потерять, пользуясь свободным от распределения критерием, когда предположения нормальной теории в действительности справедливы (хотя, конечно, мы и не знали бы этого на практике). Если эти потери малы, то это поощряет нас пожертвовать небольшой добавочной эффективностью методов стандартной нормальной теории ради расширения области применимости, связанного с использованием свободного от распределения критерия.

Мы можем сделать еще один шаг в этом сравнении критериев нормальной теории с критериями, свободными от распределения. В некоторых случаях имеется возможность исследовать относительную эффективность двух методов для широкого класса исходных распределений; и надо особенно отметить, что нет никаких оснований ожидать, что метод нормальной теории будет сохранять свое преимущество в эффективности над свободным от распределения методом, когда исходное распределение в действительности не нормально. В самом деле, можно даже предположить, что свободные от распределения методы меньше пострадают от отсутствия нормальности, чем методы нормальной теории, основанные на этом предположении. Те немногочисленные исследования, которые были проведены до сих пор, по-видимому, в целом подтверждают это предположение.

Критерии независимости

31.18 Подробное рассмотрение свободных от распределения критериев для непараметрических гипотез, которое будет иллюстрировать общие положения из **31.15—17**, мы начинаем с задачи (4) из **31.13** — задачи независимости.

Предположим, что у нас имеется выборка из непрерывной двумерной функции распределения $F(x, y)$ с непрерывными маргинальными функциями распределения $G(x)$, $H(y)$, состоящая из n пар (x, y). Мы хотим проверить гипотезу

$$H_0\colon\ F(x, y) = G(x)\,H(y) \text{ при всех } x, y. \qquad (31.9)$$

При гипотезе H_0 все $n!$ возможных упорядочений значений x

равновероятны, и независимо от x то же самое справедливо относительно $n!$ упорядочений y; мы, следовательно, имеем $(n!)^2$ равновероятных точек в пространстве выборок. Однако поскольку вопрос касается связи между x и y, для нас играют роль только различные способы соединения в пары n иксов и n игреков, и имеются $n!$ различных множеств таких соединений (получаемых, например, если фиксировать значения y и брать перестановки значений x), имеющих равные вероятности $(n!)^{-1}$. Согласно **31.16** всякий подобный критерий размера α для H_0 содержит $\alpha n! = N$ способов соединения в пары (предполагается, что N — целое положительное).

Каждое из $n!$ множеств таких пар содержит n значений (x, y) (некоторые из которых могут, конечно, совпадать). Теперь спрашивается: какую функцию от (x, y) взять в качестве статистики критерия?

Рассмотрим альтернативную гипотезу H_1, состоящую в том, что x и y имеют двумерное нормальное распределение с ненулевым коэффициентом корреляции ρ. Согласно (16.47) и (16.50) мы можем тогда записать функцию правдоподобия в виде

$$L(x|\,H_1) = \{2\pi\sigma_x\sigma_y(1-\rho^2)^{1/2}\}^{-n} \exp\left\{-\frac{n}{2(1-\rho^2)}\left[\left(\frac{\bar{x}-\mu_x}{\sigma_x}\right)^2 - \right.\right.$$

$$\left.\left. -2\rho\left(\frac{\bar{x}-\mu_x}{\sigma_x}\right)\left(\frac{\bar{y}-\mu_y}{\sigma_y}\right) + \left(\frac{\bar{y}-\mu_y}{\sigma_y}\right)^2 + \left(\frac{s_x^2}{\sigma_x^2} - \frac{2\rho r s_x s_y}{\sigma_x\sigma_y} + \frac{s_y^2}{\sigma_y^2}\right)\right]\right\}. \quad (31.10)$$

Различные способы соединения x и y в пары не влияют на наблюденные средние и дисперсии $\bar{x}$, $\bar{y}$, s_x^2, s_y^2, но они воздействуют на выборочный коэффициент корреляции r через член $\sum_{i=1}^{n} x_i y_i$ в его числителе. Очевидно, наибольшее значение (31.10) будет достигаться при любом $\rho > 0$, когда r сколь возможно велико, и при любом $\rho < 0$, когда r сколь возможно мало. По лемме Неймана — Пирсона из **22.10** мы получаем наиболее мощный критерий перестановок, выбирая в качестве нашей критической области множество тех соединений в пары, которые максимизируют (31.10), так как при гипотезе H_0 все способы образования пар равновероятны. Таким образом, рассмотрение нормальных альтернатив приводит к следующему критерию, впервые предложенному из интуитивных соображений Питмэном (1937b): отвергнуть H_0 против альтернатив положительной корреляции, если r велико, против альтернатив отрицательной корреляции, если r мало, и против общих альтернатив отсутствия независимости, если $|r|$ велико. Критическое значение в каждом случае должно определяться по распределению r среди $n!$ различных множеств пар, равновероятных при гипотезе H_0.

Хотя корреляционный критерий Питмэна дает наиболее мощный критерий перестановок для гипотезы независимости против нормальных альтернатив, он, конечно, остается в силе (т. е. имеет точный размер α) при любых альтернативах, и можно предполагать, что он будет иметь достаточно хорошую мощность в широком классе альтернатив, близких к нормальным.

Перестановочное распределение r

31.19 Поскольку

$$r = \left(\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}\right) \Big/ s_x s_y, \tag{31.11}$$

только $\sum_i x_i y_i$ является случайной величиной при перестановках. Мы можем получить ее точное распределение, а следовательно и распределение r, путем перебора всех $n!$ возможностей, но это становится слишком трудоемким на практике, когда n велико. Вместо этого мы найдем аппроксимацию к точному распределению, подбирая распределение по его моментам. Сохраняя y_i фиксированными и переставляя значения x, мы находим

$$\mathsf{M}\left(\sum x_i y_i\right) = \sum y_i \,\mathsf{M}(x_i) = \sum y_i \bar{x} = n\bar{x}\bar{y},$$

откуда согласно (31.11)

$$\mathsf{M}(r) = 0. \tag{31.12}$$

Для удобства мы примем теперь за начало отсчета средние $(\bar{x}, \bar{y})$. Имеем

$$\mathsf{D}\left(\sum_i x_i y_i\right) = \sum_i y_i^2 \,\mathsf{D}x_i + \sum_{i\neq j}\sum y_i y_j \operatorname{cov}(x_i, x_j) =$$

$$= \sum_i y_i^2 s_x^2 + \sum_{i\neq j}\sum y_i y_j \cdot \frac{1}{n(n-1)} \sum_{i\neq j}\sum x_i x_j =$$

$$= n s_y^2 s_x^2 + \left\{\left(\sum_i y_i\right)^2 - \sum_i y_i^2\right\} \frac{1}{n(n-1)} \left\{\left(\sum_i x_i\right)^2 - \sum_i x_i^2\right\} =$$

$$= n s_y^2 s_x^2 + n s_y^2 s_x^2/(n-1) = n^2 s_y^2 s_x^2/(n-1).$$

Поэтому (31.11) дает

$$\mathsf{D}(r) = \left(n^2 s_x^2 s_y^2\right)^{-1} \mathsf{D}\left(\sum xy\right) = 1/(n-1). \tag{31.13}$$

Первые два момента r, даваемые формулами (31.12) и (31.13), совсем не зависят от наблюденных значений (x, y).

Аналогичными методами находим, что

$$\left.\begin{aligned} \mathbf{M}(r^3) &= \frac{n-2}{n(n-1)^2}\left(\frac{k_3}{k_2^{3/2}}\right)\left(\frac{k_3'}{(k_2')^{3/2}}\right), \\ \mathbf{M}(r^4) &= \frac{3}{n^2-1}\left\{1+\frac{(n-2)(n-3)}{3n(n-1)^2}\left(\frac{k_4}{k_2^2}\right)\left(\frac{k_4'}{(k_2')^2}\right)\right\}, \end{aligned}\right\} \tag{31.14}$$

где k обозначают k-статистики наблюденных иксов, а k' обозначают k-статистики наблюденных игреков. Пренебрегая различием между k-статистиками и выборочными семиинвариантами, мы можем переписать (31.14) в виде

$$\left.\begin{aligned} \mathbf{M}(r^3) &\approx \frac{n-2}{n(n-1)^2}\,g_1 g_1', \\ \mathbf{M}(r^4) &\approx \frac{3}{n^2-1}\left\{1+\frac{(n-2)(n-3)}{3n(n-1)^2}\,g_2 g_2'\right\}, \end{aligned}\right\} \tag{31.15}$$

где g_1, g_2 — меры асимметрии и эксцесса для иксов, а g_1', g_2' — то же для игреков. Если они фиксированы, то (31.14) можно записать как

$$\mathbf{M}(r^3) = O(n^{-2}), \qquad \mathbf{M}(r^4) = \frac{3}{n^2-1}\{1+O(n^{-1})\}. \tag{31.16}$$

Таким образом, при $n \to \infty$ мы имеем приближенно

$$\mathbf{M}(r^3) = 0, \qquad \mathbf{M}(r^4) = \frac{3}{n^2-1}. \tag{31.17}$$

Моменты (31.12), (31.13) и (31.17) в точности равны моментам распределения (16.62) — симметричного точного распределения r в выборках из двумерного нормального распределения с $\rho = 0$, что легко можно проверить интегрированием r^2 и r^4 в (16.62). Таким образом, с хорошей степенью приближения перестановочное распределение r дается тем же выражением

$$dF = \frac{1}{\mathrm{B}\{(n-2)/2,\ 1/2\}}(1-r^2)^{(n-4)/2}\,dr, \qquad -1 \leqslant r \leqslant 1, \tag{31.18}$$

и мы можем, следовательно, использовать (31.18) или эквивалентный факт, что $t = \{(n-2)r^2/(1-r^2)\}^{1/2}$ имеет распределение Стьюдента с $n-2$ степенями свободы, для реализации наших критериев над r. В действительности (31.18) дает очень хорошее приближение даже при малых n, как можно было ожидать из точного совпадения его первых двух моментов с моментами перестановочного распределения.

Сходимость перестановочного распределения и распределений нормальной теории к одному и тому же предельному нормальному распределению была строго доказана Гёфдингом (1952).

31.20 Может показаться неожиданным, что перестановочное распределение r, используемое при проверке непараметрической гипотезы (31.9), так близко согласуется с точным распределением (16.62), которое было получено в предположении независимости и нормальности x и y. Но следует обратить внимание на то, что степень аппроксимации третьего и четвертого моментов перестановочного распределения r зависит от значений g в (31.15): они становятся малыми, когда $F(x, y)$ близко к нормальному. На самом деле мы сейчас наблюдаем, так сказать с другого конца, явление, о котором говорилось в **31.9**, а именно устойчивость распределения r, когда $\rho = 0$.

Но если фактическое совпадение перестановочного распределения с распределением нормальной теории не является полной неожиданностью, то оно во всяком случае очень удобно и приятно, так как мы можем по-прежнему пользоваться таблицами нормальной теории (в данном случае t-распределения Стьюдента) для свободного от распределения критерия непараметрической гипотезы независимости.

Ранговые критерии независимости

31.21 Некоторое неудобство r как критерия независимости, о котором кратко было сказано после (31.11), состоит в том, что вычисление его точного распределения при малых значениях n (скажем, от 5 до 10) очень трудоемко. Причина этого в том, что точное распределение r зависит от наблюденных значений (x, y), а они представляют собой, конечно, случайные величины. Несмотря на отличное приближение распределения распределением **31.18**, представляет интерес выяснить, как эта трудность может быть преодолена: это может быть полезно и в других ситуациях, так как приближение к перестановочному распределению не всегда бывает таким хорошим.

Наиболее очевидное средство устранения зависимости перестановочного распределения от случайно варьирующих наблюдений состоит в замене значений (x, y) новыми значениями (X, Y) (с коэффициентом корреляции R), определенными так, чтобы перестановочное распределение R было одним и тем же для каждой выборки (хотя, конечно, само R будет меняться от выборки к выборке). Мы, таким образом, ищем набор условных чисел (X, Y) для замены наблюденных (x, y). Как их следует выбрать? (X, Y) не должны зависеть от фактических значений (x, y), но они, очевидно, должны отражать порядковые соотношения между наблюденными значениями x и y, так как нас интересует зависимость между этими случайными величинами. Мы, таким образом, приходим к рассмотрению функций от

рангов x и y. Мы определяем ранг наблюдения y_i как его положение среди порядковых статистик, т. е.

$$\text{ранг}\{y_{(i)}\} = i.$$

Наше стремление рассматривать критерии, основанные на рангах (называемые иначе «ранговыми критериями»), поддерживается тем фактом, что ранги инвариантны при любых монотонных преобразованиях переменных. Любое такое преобразование оставляет инвариантной также гипотезу независимости (31.9), и поэтому естественно использование рангов. Мы еще не определили, какие функции от рангов использовать в качестве чисел (X, Y)[2]; самая простая и очевидная процедура — использовать сами ранги, т. е. заменить наблюденные значения x их рангами среди иксов, а наблюденные значения y — своими рангами.

31.22 Действуя таким образом, мы находим коэффициент корреляции R между n парами (X, Y), где $(X_1, X_2, \ldots, X_n)$ — перестановка первых n натуральных чисел, а $(Y_1, Y_2, \ldots, Y_n)$ — другая такая перестановка. Для получения перестановочного распределения R мы можем, как и прежде, зафиксировать значения Y и переставлять значения X, так как имеется только $n!$ различных и равновероятных наборов попарных соединений (X, Y). Поэтому мы можем без потери общности упорядочить n пар в любой выборке так, чтобы ранги Y были расположены в естественном порядке $1, 2, \ldots, n$. Если ранг X, соответствующий значению $Y = i$, обозначить X_i, то мы получим выражение для *рангового коэффициента корреляции*

$$R = \left[\frac{1}{n}\sum_{i=1}^{n} iX_i - \left\{\frac{1}{2}(n+1)\right\}^2\right] \Big/ \left\{\frac{1}{12}(n^2-1)\right\}, \qquad (31.19)$$

так как среднее и дисперсия первых n натуральных чисел равны соответственно $\frac{1}{2}(n+1)$ и $\frac{1}{12}(n^2-1)$. Обычно R называют ранговым коэффициентом корреляции Спирмена по имени знаменитого психолога, который ввел его более пятидесяти лет назад в качестве замены для обычного коэффициента корреляции; его обозначают обычно символом r_s, которым мы теперь и будем пользоваться. Так как

$$\sum_{i=1}^{n} iX_i \equiv \frac{1}{6}n(n+1)(2n+1) - \frac{1}{2}\sum_{i=1}^{n}(X_i - i)^2, \qquad (31.20)$$

то r_s можно иначе определить формулой

$$r_s = 1 - \frac{6}{n(n^2-1)} \sum_{i=1}^{n} (X_i - i)^2, \qquad (31.21)$$

которая обычно более удобна для вычислений.

31.23 Так как формулы (31.12)—(31.14) для точных моментов r справедливы для произвольных x и y, они справедливы, в частности, для r_s, определяемого формулой (31.21). Кроме того, натуральные числа имеют вследствие симметрии нулевые нечетные моменты относительно среднего. Отсюда следует, что точное распределение r_s симметрично и, следовательно, его нечетные моменты равны нулю. Если мы подставим также значения k_4, k_4' в (31.14), мы получим для точных моментов выражения

$$\left.\begin{aligned} &\mathbf{M}(r_s) = \mathbf{M}(r_s^3) = 0, \\ &\mathbf{D}(r_s) = \frac{1}{n-1}, \\ &\mathbf{M}(r_s^4) = \frac{3}{n^2-1}\left\{1 + \frac{12(n-2)(n-3)}{25n(n-1)^2}\right\}. \end{aligned}\right\} \qquad (31.22)$$

Однако, как показывают рассуждения пункта **31.21**, точное распределение r_s может быть на самом деле табулировано раз навсегда. Кендалл (1962) дает таблицы распределения $\sum_i (X_i - i)^2$, случайной составляющей r_s в (31.21), для $n = 4(1)10$. («Хвосты» этих распределений воспроизведены в *Biometrika Tables* *).) За этими пределами приближение (31.18) вполне удовлетворительно для практических целей, что показывает следующая таблица, в которой сравниваются точные и приближенные критические значения r_s для размеров критерия $\alpha = 0{,}05$, $0{,}01$ и $n = 10$.

Сравнение точных и приближенных критических значений r_s при n=10

Двусторонний критерий	Точные критические значения (из Кендалла (1962))	Приближенные критические значения из (31.18)
$\alpha = 0{,}05$	0,648	0,632
$\alpha = 0{,}01$	0,794	0,765

*) Таблица из *Biometrika Tables* приведена в *Таблицах математической статистики*. Полностью точное распределение $\sum (X_i - i)^2$ для $n = 2(1)11$ дано в *Сборнике статистических таблиц*. (*Прим. ред.*)

31.24 Мы выбрали r_s среди возможных ранговых критериев независимости по причине его простоты; ясно, что любая разумная мера корреляции между x и y, основанная на их рангах, дает критерий независимости. Даниэлс (1944) определил класс коэффициентов корреляции, включающий как обычный, так и ранговые коэффициенты корреляции, и в дальнейшем показал (Даниэлс, 1948), что все они являются, по существу, *коэффициентами беспорядка* в том смысле, что если два значения y поменять местами так, чтобы они стали расположены в том же порядке, что и соответствующие значения x, то величина любого коэффициента из этого класса увеличится. Мы рассмотрим вопрос об измерении беспорядка среди рангов x и y.

Предположим, как в **31.22**, что ранги y (которые там обозначались Y) расположены в натуральном порядке $1, 2, \ldots, n$ и что соответствующие ранги x, образующие перестановку чисел $1, 2, \ldots, n$, равны $X_1, X_2, \ldots, X_n$. Естественный метод измерения беспорядка x-рангов, т. е. отклонения от порядка $1, 2, \ldots, n$, состоит в подсчете числа инверсий между ними. Например, при $n = 4$ в x-ранжировке 3214 имеются три инверсии, а именно 3—2, 3—1, 2—1. Число таких инверсий, которое мы будем обозначать Q, может изменяться в пределах от 0 до $\frac{1}{2}n(n-1)$, причем эти граничные значения достигаются соответственно на x-ранжировках $1, 2, \ldots, n$ и $n, (n-1), \ldots, 1$. Мы можем, следовательно, определить коэффициент

$$t = 1 - \frac{4Q}{n(n-1)}, \tag{31.23}$$

который распределен симметрично на отрезке $(-1, +1)$ над множеством из $n!$ равновероятных перестановок и, следовательно, имеет математическое ожидание 0, когда выполняется (31.9).

Коэффициент (31.23) рассматривался несколькими авторами (Фехнер, Липпс) в период около 1900 г. и позже рядом авторов, особенно Линдебергом, в 20-х годах (исторические подробности даются Крускалом (1958)), но впервые получил широкое применение после серии работ М. Кендалла, начинающейся с 1938 г. и собранной в монографии (Кендалл, 1962), к которой следует также обращаться по вопросам использования t и r_s как *мер* корреляции. Здесь мы рассматриваем их только в качестве свободных от распределения критериев для гипотезы (31.9).

31.25 Распределение коэффициента t, или, что эквивалентно, числа инверсий Q над множеством $n!$ равновероятных x-ранжировок легко установить с помощью производящих функций. Пусть $f(Q, n)/n!$ — функция частот для Q при объеме выборки n. Мы можем произвести $n!$ x-ранжировок для объема выборки n из $(n-1)!$ ранжировок для объема выборки $(n-1)$, помещая новый ранг «n» в любое возможное положение по отноше-

нию к имеющимся $(n-1)$. (Таким образом, например, 2! ранжировок для $n=2$

$$\begin{matrix}1 & 2\\ 2 & 1\end{matrix}$$

переходят в 3! ранжировок для $n=3$

$$\begin{matrix}312 & 132 & 123\\ 321 & 231 & 213.\end{matrix}$$

В любой ранжировке увеличение Q, вызванное этим процессом, в точности равно числу рангов справа от точки, в которой помещается «n». Любое значение Q в n-ранжировке, таким образом, представляет собой сумму n членов, соответствующих различным значениям Q в $(n-1)$-ранжировке. Это приводит к соотношению

$$f(Q,\ n)=f(Q,\ n-1)+f(Q-1,\ n-1)+f(Q-2,\ n-1)+\ \ldots$$
$$\ldots\ +f(Q-(n-1),\ n-1). \qquad (31.24)$$

Если теперь $f(Q, n)$ есть коэффициент при θ^Q в производящей функции $G(\theta, n)$, то из (31.24) следует, что

$$G(\theta,\ n)=G(\theta,\ n-1)+\theta G(\theta,\ n-1)+\theta^2 G(\theta,\ n-1)+\ \ldots$$
$$\ldots\ +\theta^{n-1}G(\theta,\ n-1)=\frac{\theta^n-1}{\theta-1}\,G(\theta,\ n-1). \qquad (31.25)$$

Повторным применением (31.25) мы находим

$$G(\theta,\ n)=\left(\frac{\theta^n-1}{\theta-1}\right)\left(\frac{\theta^{n-1}-1}{\theta-1}\right)\ldots\left(\frac{\theta^3-1}{\theta-1}\right)G(\theta,\ 2), \qquad (31.26)$$

и так как мы непосредственно видим, что

$$G(\theta,\ 2)=1\cdot\theta^0+1\cdot\theta^1=\frac{\theta^2-1}{\theta-1},$$

то (31.26) можно переписать как

$$G(\theta,\ n)=\prod_{s=1}^{n}\left(\frac{\theta^s-1}{\theta-1}\right). \qquad (31.27)$$

Мы получим характеристическую функцию Q, добавляя множитель $(n!)^{-1}$ и заменяя θ на $\exp(i\theta)$ в (31.27), так что

$$\varphi(\theta)=\{n!\,(e^{i\theta}-1)^n\}^{-1}\prod_{s=1}^{n}(e^{i\theta s}-1). \qquad (31.28)$$

Следовательно, п. ф. с. Q равна

$$\psi(\theta)=\sum_{s=1}^{n}\log(e^{i\theta s}-1)-n\log(e^{i\theta}-1)-\log(n!). \qquad (31.29)$$

Если мы всюду в (31.29) подставим

$$e^{i\theta s}-1\equiv e^{i\theta s/2}2\,\mathrm{sh}\left(\frac{1}{2}i\theta s\right),$$

то (31.29) перейдет в

$$\psi(\theta)=$$

$$=\frac{1}{2}i\theta\left(\sum_{s=1}^{n}s-n\right)+\sum_{s=1}^{n}\log\mathrm{sh}\left(\frac{1}{2}i\theta s\right)-n\log\mathrm{sh}\left(\frac{1}{2}i\theta\right)-\log(n!)=$$

$$=\frac{1}{4}n(n-1)i\theta+\sum_{s=1}^{n}\log\left(\frac{\mathrm{sh}\left(\frac{1}{2}i\theta s\right)}{\frac{1}{2}i\theta s}\right)-n\log\left(\frac{\mathrm{sh}\left(\frac{1}{2}i\theta\right)}{\frac{1}{2}i\theta}\right);\qquad(31.30)$$

пользуясь (3.61), преобразуем (31.30) к виду

$$\psi(\theta)=\frac{1}{4}n(n-1)i\theta+\sum_{j=1}^{\infty}\frac{B_{2j}(i\theta)^{2j}}{2j(2j)!}\sum_{s=1}^{n}(s^{2j}-1),\qquad(31.31)$$

где B_{2j} — (ненулевые) числа Бернулли четного порядка, определенные в **3.25**.

Беря коэффициенты при $(i\theta)^{2j}/(2j)!$ в (31.31), мы получаем для семиинвариантов Q

$$\left.\begin{aligned}&\varkappa_1=\frac{1}{4}n(n-1),\quad \varkappa_{2j+1}=0,\quad j\geqslant 1,\\&\varkappa_{2j}=\frac{B_{2j}}{2j}\left(\sum_{s=1}^{n}s^{2j}-n\right).\end{aligned}\right\}\qquad(31.32)$$

Вследствие (31.23) мы получаем отсюда семиинварианты для самого коэффициента ранговой корреляции t:

$$\left.\begin{aligned}&\varkappa_{2j+1}=0,\quad j\geqslant 0,\\&\varkappa_{2j}=\frac{2^{4j-1}B_{2j}}{j\{n(n-1)\}^{2j}}\left(\sum_{s=1}^{n}s^{2j}-n\right).\end{aligned}\right\}\qquad(31.33)$$

Таким образом, t имеет распределение, симметричное относительно нуля, и

$$\mathrm{D}t=\frac{2^3B_2}{\{n(n-1)\}^2}\left\{\frac{1}{6}n(n+1)(2n+1)-n\right\}=\frac{2(2n+5)}{9n(n-1)}.\qquad(31.34)$$

31.26 Далее, (31.33) показывает, что $\varkappa_{2j}$ имеет по n порядок $n^{-4j}\sum_{s=1}^{n} s^{2j}$. Так как сумма имеет порядок n^{2j+1}, это значит, что

$$\varkappa_{2j} = O(n^{1-2j}),$$

и, следовательно, нормированные семиинварианты равны

$$\frac{\varkappa_{2j}}{(\varkappa_2)^j} = O(n^{1-j}).$$

Таким образом,

$$\lim_{n\to\infty} \frac{\varkappa_{2j}}{(\varkappa_2)^j} = 0, \qquad j > 1, \tag{31.35}$$

и, следовательно, распределение t стремится к нормальному с нулевым средним и дисперсией, определяемой (31.34). Сходимость к нормальному распределению чрезвычайно быстрая. Кендалл (1962) приводит точную функцию распределения (полученную из (31.24)) для $n = 4(1)10$. За этими пределами асимптотическое нормальное распределение может быть использовано с малой потерей в точности *).

31.27 В **31.24** мы пришли к коэффициенту t из тех соображений, что число инверсий Q является естественной мерой беспорядка в x-ранжировке. При дальнейшем размышлении кажется разумным придавать инверсиям различный вес; например, в x-ранжировке 24351 чувствуется, что инверсия 5—1 должна иметь больший вес, чем инверсия 4—3, поскольку она представляет более серьезное отклонение от натурального порядка 1, 2, ..., n. Сам собой напрашивается простой способ взвешивания — посредством расстояния между рангами, образующими инверсию; в приведенном только что примере это дало бы соответственно веса 4 и 1 двум инверсиям. Таким образом, если мы положим

$$h_{ij} = \begin{cases} +1, & \text{если } X_i > X_j, \\ 0 & \text{в противном случае,} \end{cases} \tag{31.36}$$

то мы теперь хотим использовать взвешенную сумму инверсий

$$V = \sum_{i<j}\sum h_{ij}(j-i) \tag{31.37}$$

вместо нашей предыдущей суммы инверсий

$$Q = \sum_{i<j}\sum h_{ij}. \tag{31.38}$$

*) Эта таблица воспроизведена в *Таблицах математической статистики*. Точное распределение статистики Q для $n = 2(1)12$ дается в *Сборнике статистических таблиц*. (*Прим. ред.*)

Однако использование (31.37) возвращает нас снова к r_s. Мы оставляем читателю доказать в упражнении 31.5, что

$$V \equiv \frac{1}{2}\sum_{i=1}^{n}(X_i - i)^2, \tag{31.39}$$

так что согласно (31.21)

$$r_s = 1 - \frac{12V}{n(n^2-1)}, \tag{31.40}$$

что дает определение r_s, аналогичное определению (31.23) для t.

37.28 Имеет место замечательный факт, состоящий в том, что, несмотря на использование весьма различных методов взвешивания инверсий, Q и V, определяемые из (31.37), (31.38), а следовательно, и статистики t и r_s очень сильно коррелированы, когда выполняется гипотеза независимости (31.9), — мы предоставляем читателю в упражнении 31.6 получить настоящее значение их коэффициента корреляции. Он убывает от 1 при $n = 2$ (когда t и r_s эквивалентны) до своего минимального значения 0,98 при $n = 5$ и затем возрастает к 1 при $n \to \infty$. Таким образом, эти критерии асимптотически эквивалентны, когда справедлива гипотеза H_0, а это вместе с результатом **25.13** означает, что с точки зрения асимптотической относительной эффективности оба критерия обладают одинаковыми свойствами. Даниэлс (1944) показал, что предельное совместное распределение t и r_s, когда справедлива H_0, есть двумерное нормальное распределение.

31.29 В выборках из двумерной нормальной генеральной совокупности высокая корреляция между t и r_s сохраняется, даже когда коэффициент корреляции исходного распределения $\rho \neq 0$; С. Дэвид и др. (1951) показали, что при $n \to \infty$ t и r_s имеют коэффициент корреляции, стремящийся к значению $\geqslant 0{,}984$, если $|\rho| \leqslant 0{,}8$, и к 0,937, если $\rho = 0{,}9$.

Гёфдинг (1948а) показал, что t и r_s в самом общем случае имеют асимптотическое двумерное нормальное распределение, но что их коэффициент корреляции сильно зависит от исходного двумерного распределения и может в действительности равняться нулю.

Эффективности критериев независимости

31.30 Мы теперь исследуем асимптотические относительные эффективности (АОЭ) трех рассмотренных критериев независимости по отношению к обычному выборочному коэффициенту корреляции r, когда альтернативной гипотезой служит двумерное нормальное распределение, как в (31.10). С помощью методов **23.27—36** мы видим, что r дает РНМН критерий для $\rho = 0$ против односторонних и двусторонних альтернатив, — читателю предлагается проверить это в упражнении 31.21. Так как со-

гласно **31.19** перестановочный критерий, основанный на r, асимптотически эквивалентен r-критерию нормальной теории для независимости, мы видим, что его АОЭ по сравнению с этим критерием равна 1.

31.31 Мы теперь выведем АОЭ критерия, основанного на статистике t, определенной в (31.23). По определению (31.36) мы видим, что

$$h_{ij} = \tfrac{1}{2}\{1 - \operatorname{sign}(x_i - x_j)\operatorname{sign}(y_i - y_j)\},$$

и, так как в сумме $Q = \sum\sum_{i<j} h_{ij}$ имеются $\frac{1}{2}n(n-1)$ членов, мы получаем для их среднего

$$\mathbf{M}\left\{\frac{Q}{\frac{1}{2}n(n-1)}\right\} = \mathbf{M}(h_{ij}) =$$

$$= \tfrac{1}{2}\{1 - \mathbf{M}[\operatorname{sign}(x_i - x_j)\operatorname{sign}(y_i - y_j)]\}, \qquad (31.41)$$

что вместе с (31.23) дает

$$\mathbf{M}(t) = 1 - 2\mathbf{M}\left\{\frac{Q}{\frac{1}{2}n(n-1)}\right\} = \mathbf{M}[\operatorname{sign}(x_i - x_j)\operatorname{sign}(y_i - y_j)].$$

Если x и y имеют двумерное нормальное распределение F с коэффициентом корреляции ρ, то такое же утверждение верно и для $w = (x_i - x_j)$ и $z = (y_i - y_j)$. Таким образом,

$$\mathbf{M}(t) = \mathbf{M}[\operatorname{sign}(x_i - x_j)\operatorname{sign}(y_i - y_j)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \operatorname{sign} w \operatorname{sign} z \, dF,$$

и с помощью (4.8) это приводится к виду

$$= \int_{-\infty}^{\infty}\left\{\frac{1}{\pi^2}\int_{-\infty}^{\infty}\frac{\sin t_1 w}{t_1}\,dt_1 \int_{-\infty}^{\infty}\frac{\sin t_2 z}{t_2}\,dt_2\right\} dF,$$

что можно переписать как

$$= \frac{1}{\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp(it_1 w + it_2 z)\,dF\right\}\frac{dt_1}{it_1}\frac{dt_2}{it_2}. \qquad (31.42)$$

Внутренний двойной интеграл в (31.42) представляет собой х. ф. F, равную

$$\varphi(t_1, t_2) = \exp\left\{-\tfrac{1}{2}(t_1^2 + t_2^2 + 2\rho t_1 t_2)\right\}.$$

Подставляя это выражение и дифференцируя оставшийся двойной интеграл по ρ, мы находим

$$\frac{\partial}{\partial\rho}\mathsf{M}(t)=\frac{1}{\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(t_1,\,t_2)\,dt_1\,dt_2. \tag{31.43}$$

Но двойной интеграл в правой части (31.43) легко находится:

$$\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2}t_2^2(1-\rho^2)\right\}\left[\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2}(t_1+\rho t_2)^2\right\}dt_1\right]dt_2=\frac{2\pi}{(1-\rho^2)^{1/2}}.$$

Таким образом, (31.43) принимает вид

$$\frac{\partial}{\partial\rho}\mathsf{M}(t)=\frac{2}{\pi(1-\rho^2)^{1/2}},$$

так что

$$\left[\frac{\partial}{\partial\rho}\mathsf{M}(t)\right]_{\rho=0}=\frac{2}{\pi}. \tag{31.44}$$

Кроме того, из (31.34)

$$\mathsf{D}(t\,|\,H_0)\sim\frac{4}{9n}, \tag{31.45}$$

тогда как для обычного коэффициента корреляции r согласно (26.31)

$$\left[\frac{\partial}{\partial\rho}\mathsf{M}(r)\right]_{\rho=0}\to 1, \tag{31.46}$$

а из **31.19**

$$\mathsf{D}(r)\sim\frac{1}{n}. \tag{31.47}$$

Применяя (25.27) с $m=1$, $\delta=1/2$, из результатов (31.44)—(31.47) получаем для АОЭ t по сравнению с r

$$A_{t,\,r}=9/\pi^2. \tag{31.48}$$

Согласно замечанию **31.28** (31.48) будет справедливо также для АОЭ r_s по сравнению с r — результат, принадлежащий первоначально Хотеллингу и Пэбсту (1936).

31.32 Помимо результатов пунктов **31.30—31**, относящихся к двумерным нормальным альтернативам, имеется мало работ об эффективности критериев независимости, что связано в основном с трудностью задания альтернатив к независимости при отсутствии нормальности. Достойным внимания исключением является работа Конейна (1956), в которой рассматривается класс альтернатив к независимости, порождаемых линейными преобразованиями двух независимых величин. Он находит, как и выше, что критерии t и r_s часто асимптотически эквивалентны и каждый из них по АОЭ близок

к критерию, основанному на выборочном коэффициенте корреляции r, и может даже (в случае двойного экспоненциального распределения) превосходить его.

31.33 Недостатком всех критериев, которые мы рассматривали, является то, что они не обладают состоятельностью против *любых* отклонений от гипотезы независимости (31.9). Чтобы понять это, достаточно заметить, что каждый из них основан, по существу, на том или ином коэффициенте корреляции, распределение которого не зависит от параметров сдвига и масштаба, но будет зависеть от теоретического коэффициента корреляции ρ. При отклонениях от независимости, для которых $\rho \neq 0$, эти критерии будут состоятельными. Но для распределений, отличных от нормальных, вполне возможны ситуации, когда при отсутствии независимости $\rho = 0$ (см. **26.6**), и мы не можем ожидать от наших критериев состоятельности против таких альтернатив. Имея это в виду, Гёфдинг (1948b) предложил другой свободный от распределения критерий для гипотезы (31.9), который состоятелен против любого непрерывного альтернативного двумерного распределения с непрерывными маргинальными распределениями. Гёфдинг табулировал распределение своей статистики для $n = 5, 6, 7$ и получил ее предельную х. ф. и ее семиинварианты. (Предельная ф. р. дается в работе Блюма и др. (1961).) Он доказал также, что для этого класса альтернатив не существует рангового критерия независимости, который был бы несмещенным для любого размера $\alpha = M/n!$. Однако если допускается рандомизация в критической функции, то, как показал Леман (1951), несмещенные ранговые критерии независимости, вообще говоря, существуют.

Критерии случайности против альтернатив тренда

31.34 Как мы отмечали в **31.14**, проблема (3) из **31.13**, состоящая в проверке гипотезы

$$H_0\colon F_1(x) = F_2(x) = \ \dots\ = F_n(x) \text{ при всех } x, \qquad (31.49)$$

когда имеется по одному наблюдению из каждого из n непрерывных распределений, упорядоченных согласно значениям некоторой переменной y, эквивалентна проверке независимости x и y. Поэтому любой из наших критериев независимости может быть использован как критерий случайности. Однако, поскольку переменная y обычно не является случайной величиной, а только упорядочивает распределения (по времени или как-нибудь иначе), любое монотонное преобразование y играло бы ту же роль, что и само y. Поэтому естественно ограничиться ранговыми критериями случайности, так как ранги инвариантны при монотонном преобразовании, оставляющем гипотезу (31.49) без изменения.

Манн (1945) был, по-видимому, первым, кто понял, что статистика ранговой корреляции может быть использована не только для проверки независимости, но и случайности, и предложил использовать t (хотя, конечно, с таким же успехом можно было бы использовать r_s) против класса альтернатив

$$H_1\colon F_1(x) < F_2(x) < \ \dots\ < F_n(x) \text{ при всех } x, \qquad (31.50)$$

где наблюдения x_i остаются независимыми.

Так как (31.50) утверждает, что вероятность наблюдению попасть ниже любого фиксированного значения монотонно возрастает, когда мы продвигаемся по последовательности n наблюдений, то H_1 можно назвать альтернативой *тренда вниз*. Поэтому критическая область для критерия размера α состоит из наибольших значений Q, числа инверсий, определенного в (31.38), в количестве 100α процентов от числа всех возможных значений Q.

31.35 Из (31.50) следует, что при $i < j$

$$\mathsf{P}\{h_{ij} = 1\} = \mathsf{P}\{X_i < X_j\} = \tfrac{1}{2} + \varepsilon_{ij}, \quad 0 < \varepsilon_{ij} \leqslant \tfrac{1}{2}. \tag{31.51}$$

Поэтому мы имеем согласно (31.38)

$$\mathsf{M}(Q \mid H_1) = \tfrac{1}{4} n(n-1) + \sum_{i<j} \sum \varepsilon_{ij} = \tfrac{1}{4} n(n-1) + S_n, \tag{31.52}$$

где S_n — сумма всех $\frac{1}{2} n(n-1)$ значений ε_{ij}.

Рассмотрим теперь дисперсию Q:

$$\mathsf{D}(Q \mid H_1) = \mathsf{D}\left\{ \sum_{i<j} h_{ij} \right\} = \sum_{i<j} \mathsf{D}(h_{ij}) + \sum_{i<j} \sum_{k<l} \operatorname{cov}(h_{ij}, h_{kl}). \tag{31.53}$$

Ковариационные члены в (31.53) имеются двух видов. Все члены, содержащие четыре различных индекса, равны нулю, так как случайные величины тогда независимы, и таких членов $\binom{n}{4}$. Остальные члены отличны от нуля и содержат только три различных индекса, так что (i, j) и (k, l) имеют один общий индекс. Число таких членов имеет порядок $\binom{n}{3}$ — числа способов выбрать три индекса из n. Следовательно, поскольку в первой сумме в (31.53) только $\binom{n}{2}$ членов, мы можем написать

$$\mathsf{D}(Q \mid H_1) = O(n^3). \tag{31.54}$$

Из **31.26** следует, что статистика Q при гипотезе H_0 распределена асимптотически нормально, и, таким образом, критическая область критерия состоит асимптотически из значений Q, превосходящих значение

$$Q_0 = \tfrac{1}{4} n(n-1) + d_\alpha \left\{ \tfrac{1}{72} n(n-1)(2n+5) \right\}^{1/2}, \tag{31.55}$$

где член в фигурных скобках есть дисперсия Q (полученная из (31.34) и (31.23)), а d_α — соответственно выбранное нормальное отклонение.

31.36 Из (31.52) и (31.55) мы видим, что

$$\mathbf{P}\{Q > Q_0 | H_1\} \sim \mathbf{P}\left\{ Q - \mathbf{M}(Q | H_1) > \right.$$
$$\left. > d_\alpha \left[\frac{1}{72} n(n-1)(2n+5)\right]^{1/2} - S_n \,\middle|\, H_1 \right\}. \qquad (31.56)$$

Пользуясь (31.54), мы можем асимптотически переписать (31.56) как

$$\mathbf{P}\{Q > Q_0 | H_1\} \sim \mathbf{P}\{Q - \mathbf{M}(Q | H_1) > [\mathbf{D}(Q | H_1)]^{1/2} [d_\alpha - cn^{-3/2} S_n]\}, \qquad (31.57)$$

где c — некоторая константа. Теперь мы наложим условие, что

$$n^{-3/2} S_n \to \infty \qquad (31.58)$$

при $n \to \infty$. Тогда

$$\lambda = d_\alpha - cn^{-3/2} S_n \to \infty \qquad (31.59)$$

и λ будет отрицательным, когда n достаточно велико. По неравенству Чебышева (3.94) мы имеем *a fortiori* для отрицательного λ и любой случайной величины

$$\mathbf{P}\{x - \mathbf{M}(x) > \lambda (\mathbf{D}x)^{1/2}\} \geqslant 1 - \frac{1}{\lambda^2}. \qquad (31.60)$$

Таким образом, если выполняется (31.58), то (31.57), (31.59) и (31.60) дают

$$\lim_{n \to \infty} \mathbf{P}\{Q > Q_0 | H_1\} = 1.$$

Следовательно, этот критерий случайности состоятелен при условии, что выполняется (31.58). Это довольно слабое требование, так как S_n состоит из $n(n-1)/2$ членов. Поэтому, если для ε_{ij} имеется фиксированная ненулевая нижняя граница, то (31.58) заведомо выполняется. Обычно рассматривают альтернативы, при которых ε_{ij} является функцией только от расстояния $|i-j|$; если это возрастающая функция, то (31.58) заведомо выполняется.

Помимо более общего варианта этого результата, в котором ε_{ij} не обязательно имеют один и тот же знак, Манн (1945) получил также условие несмещенности критерия, которое по существу совпадает с даваемым в упражнении 31.8.

37.37 Мы теперь рассмотрим альтернативу тренда специального вида, когда среднее величины x_i есть линейная функция от i, а ее распределение нормально с этим средним и постоянной дисперсией. Это обычная модель линейной регрессии с нормальными ошибками. Мы имеем

$$x_i = \beta_0 + \beta_1 i + \delta_i, \qquad (31.61)$$

где ошибки δ_i независимы и нормально распределены, а их дисперсия равна σ^2 при всех i. Проверка случайности эквивалентна проверке гипотезы

$$H_0\colon\ \beta_1 = 0 \tag{31.62}$$

в (31.61). Мы найдем асимптотическую относительную эффективность (АОЭ) критерия, основанного на t (или, что эквивалентно, на Q), по сравнению со стандартным критерием, основанным на выборочном коэффициенте регрессии

$$b = \frac{\sum (x_i - \bar{x})(i - \bar{i})}{\sum (i - \bar{i})^2} = \frac{\sum x_i i - \frac{1}{2} n (n+1) \bar{x}}{\frac{1}{12} n (n^2 - 1)}, \tag{31.63}$$

который представляет собой критерий отношения правдоподобия для (31.62) и (поскольку H_0 налагает только одно ограничение) является РНМ при односторонних альтернативах, скажем $\beta_1 < 0$, и РНМН при двусторонних альтернативах $\beta_1 \neq 0$ (см. **24.27**). Без потери общности мы полагаем $\sigma^2 = 1$. Из теории наименьших квадратов имеем (см. примеры 19.3, 19.6)

$$\mathbf{M}(b \mid H_1) = \beta_1, \qquad \mathbf{D}(b \mid H_0) = \frac{1}{\sum (i - \bar{i})^2} = \frac{12}{n (n^2 - 1)},$$

так что

$$\frac{\left\{\left[\frac{\partial \mathbf{M}(b \mid H_1)}{\partial \beta}\right]_{\beta_1 = 0}\right\}^2}{\mathbf{D}(b \mid H_0)} = \frac{n (n^2 - 1)}{12} \sim \frac{n^2}{12}. \tag{31.64}$$

31.38 Чтобы получить эквивалент (31.64) для критерия, основанного на t, нам потребуется производная от

$$\mathbf{M}(Q \mid H_1) = \mathbf{M}\left\{\sum_{i<j} h_{ij}\right\} = \sum_{i<j} \mathbf{M}(h_{ij}). \tag{31.65}$$

Согласно модели (31.61) $(x_i - x_j)$ имеет нормальное распределение со средним $\beta_1 (i - j)$ и дисперсией 2. Следовательно,

$$\mathbf{M}(h_{ij}) = \mathbf{P}\{h_{ij} = 1\} = \mathbf{P}\{x_i > x_j\} =$$

$$= \int_0^\infty \frac{1}{2\pi^{1/2}} \exp\left\{-\frac{1}{4}[t - \beta_1 (i - j)]^2\right\} dt =$$

$$= \int_{-\beta_1 (i-j)/2^{1/2}}^\infty (2\pi)^{-1/2} \exp\left(-\frac{1}{2} u^2\right) du.$$

Отсюда

$$\left[\frac{\partial}{\partial \beta_1} \mathbf{M}(h_{ij})\right]_{\beta_1 = 0} = \frac{i - j}{2^{1/2}} \cdot \frac{1}{(2\pi)^{1/2}} = \frac{(i - j)}{2\pi^{1/2}}. \tag{31.66}$$

Из (31.65) и (31.66)

$$\left[\frac{\partial}{\partial\beta_1}\mathbf{M}(Q\mid H_1)\right]_{\beta_1=0} = -\frac{1}{2\pi^{1/2}}\sum_{i<j}(j-i) =$$

$$= -\frac{1}{2\pi^{1/2}}\frac{n(n^2-1)}{6} = \frac{-n(n^2-1)}{12\pi^{1/2}}. \qquad (31.67)$$

Кроме того, из (31.34) и (31.23)

$$\mathbf{D}(Q\mid H_0) = \frac{1}{72}n(n-1)(2n+5). \qquad (31.68)$$

Из (31.67) и (31.68)

$$\frac{\left\{\left[\frac{\partial}{\partial\beta_1}\mathbf{M}(Q\mid H_1)\right]_{\beta_1=0}\right\}^2}{\mathbf{D}(Q\mid H_0)} = \frac{n^2(n^2-1)^2}{144\pi}\cdot\frac{72}{n(n-1)(2n+5)} \sim \frac{n^3}{4\pi}. \qquad (31.69)$$

Пользуясь (31.64) и (31.69), из (25.27) с $m=1$, $\delta=3/2$ получаем для АОЭ критерия Q по сравнению с b значение

$$A_{Q,b} = A_{t,b} = \left(\frac{3}{\pi}\right)^{1/3} \approx 0{,}98. \qquad (31.70)$$

Так же, как и раньше (см. **31.28**), тот же результат выполняется для коэффициента r_s (или, что эквивалентно, для V); непосредственный вывод АОЭ для V мы оставляем читателю в качестве упражнения 31.9.

Оптимальные ранговые критерии независимости и случайности

31.39 Стоит отметить, что при нормальных альтернативах ранговые коэффициенты корреляции t и r_s даже более эффективны как критерии случайности, чем как критерии двумерной независимости, поскольку значения АОЭ, даваемые (31.70) и (31.48), равны $(3/\pi)^{1/3}$ и $(3/\pi)^2$ соответственно. Но хотя оба эти значения близки к 1, они не равны 1, и остается вопрос, существуют ли для этих задач свободные от распределения критерии, для которых АОЭ по сравнению с наилучшим критерием равна 1.

Для того чтобы ответить на этот вопрос, вернемся к нашим рассуждениям в **31.21**, где r_s среди всех возможных ранговых критериев было выбрано из-за простоты. В самом деле, мы решили заменить наблюденные значения варианты x их рангами. Но поскольку АОЭ перестановочного критерия, основанного на самих значениях варианты, равна 1 при нормальных альтернативах (см. **31.30**), следует ожидать, что оптимальная эффективность сохранится, если мы заменим значения варианты функциями от их рангов, которые асимптотически полностью коррелированы со значениями варианты. Предположим теперь, что

после ранжировки наблюдений x мы заменяем их математическими ожиданиями порядковых статистик в выборке объема n из стандартного нормального распределения. Они образуют вполне определенный набор условных чисел, называемых обычно *нормальными метками* (normal scores); смысл их использования состоит в том, что при $n \to \infty$ корреляция этих чисел со значениями варианты стремится к 1, и мы получим оптимальные ранговые критерии против нормальных альтернатив. Статистика критерия, следовательно, имеет вид

$$\frac{\frac{1}{n}\sum_{i=1}^{n} iE(X_i, n) - \frac{1}{2}(n+1)\cdot\frac{1}{n}\sum_{i=1}^{n} E(X_i, n)}{\left[\frac{1}{12}(n^2-1)\cdot\frac{1}{n}\sum_{i=1}^{n}\left\{E(X_i, n) - \frac{1}{n}\sum_{i=1}^{n} E(X_i, n)\right\}^2\right]^{1/2}}, \tag{31.71}$$

где X_i теперь — ранг значения x, соответствующего i-му наибольшему значению y, а $E(s, n)$ — математическое ожидание $x_{(s)}$ в выборке объема n из стандартного нормального распределения. Пренебрегая постоянными, получаем, что (31.71) эквивалентно критерию со статистикой

$$c = \sum_{i=1}^{n} iE(X_i, n), \tag{31.72}$$

который, следовательно, имеет АОЭ, равную 1, при проверке независимости или случайности против нормальных альтернатив.

Бхучонгкул (1964) получил подтверждение этого результата, исследуя использование любых условных чисел в статистиках критериев.

Использование нормальных меток в качестве условных чисел было предложено Р. Фишером и Ф. Иэйтсом во введении к их таблицам *Statistical Tables*, впервые опубликованным в 1938 г. Свойства локальной оптимальности статистики критерия (31.72) были доказаны Гёфдингом (1950) и Терри (1952). Прямое доказательство асимптотически полной корреляции между математическими ожиданиями порядковых статистик и значениями варианты, которые они заменяют, получается из теоремы Гёфдинга (1953), утверждающей, что для любой исходной ф. р. $F(x)$ с конечным средним и любой действительной непрерывной функции $g(x)$, которая ограничена по величине интегрируемой выпуклой функцией,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} g\{E(m, n)\} = \int_{-\infty}^{\infty} g(x)\, dF. \tag{31.73}$$

Последовательная подстановка $g(x) = \cos xt$ и $g(x) = \sin xt$ в (31.73) показывает, что предельная х. ф. $E(m, n)$ совпадает с х. ф. распределения $F(x)$, равной $\mathbf{M}\{\cos xt + i \sin xt\}$.

Белл и Доксум (1965) показали, что если вместо нормальных меток $E(s, n)$ использовать наблюдения $x_{(s)}$ в некоторой выборке из n нормальных отклонений, то мы получим те же асимптотические свойства во всех ситуациях, рассмотренных в этой главе. Преимущества этого заключаются в том, что не требуется специальных таблиц и что критерии могут иметь точный размер α; недостатком является то, что при малых объемах выборок мощность этих критериев, по-видимому, ниже, чем у критериев нормальных меток. Джогдео (1966) показал, что эти критерии имеют некоторые нежелательные свойства.

Бриллинджер (1966) указал, что совокупность упорядоченных значений $x_{(s)}$ максимально коррелирована со значениями $z_{(s)} = a + b\mathbf{M}(x_{(s)})$ при любом объеме выборки. Это следует из того факта, что коэффициент корреляции между $x_{(s)}$ и $z_{(s)}$ равен корреляционному отношению x относительно своих упорядоченных значений (см. (26.40) и (26.45)).

31.40 Помимо отыскания оптимальных ранговых критериев против нормальных альтернатив, как в **31.39**, мы можем также поставить вопрос, имеются ли альтернативы, для которых данный ранговый критерий оптимален среди ранговых критериев. Мы не будем заниматься этим вопросом здесь, так как это исследование было бы искусственным с нашей теперешней точки зрения (см. **31.17**), которая, по существу, рассматривает свободные от распределения процедуры как устойчивые заменители для стандартных процедур нормальной теории. Наши интересы поэтому ограничены сравнениями эффективности между свободными от распределения методами и методами стандартной нормальной теории. Общее изложение ранговых критериев имеется в книгах Лемана (1959) и Фрэзера (1957) *).

31.41 Прежде чем закончить рассмотрение критериев случайности, заметим, что в литературе было предложено большое количество таких критериев, ни один из которых не является столь эффективным против нормальных альтернатив, как те, что мы рассматривали. Однако некоторые из них заметно проще для вычислений, чем $r_{(s)}$ и t, и лишь немногим менее эффективны. Они рассматриваются в упражнениях 31.10—31.12. Для других критериев АОЭ была получена Стьюартом (1954b, 1956).

Двухвыборочные критерии

31.42 Мы теперь рассмотрим задачу (1) из **31.13**. Имея независимые случайные выборки объемов n_1 и n_2 из непрерывных функций распределения $F_1(x)$ и $F_2(x)$ соответственно, мы хотим проверить гипотезу

$$H_0\colon\ F_1(x) = F_2(x) \quad \text{при всех} \quad x. \tag{31.74}$$

Как мы отмечали в **31.14**, это эквивалентно проверке независимости величины x и фиктивной величины y, принимающей

*) См. также книгу: Я. Гаек, З. Шидак, *Теория ранговых критериев* (перев. с англ.), М., «Наука», 1971. (*Прим. ред.*)

только два различных значения. Имеются $n_1 + n_2 = n$ наблюдений над парой (x, y).

Будем временно рассматривать n значений x размещенными по n позициям, обозначаемым

$$1,\ 2,\ 3,\ \ldots,\ n_1;\ n_1+1,\ n_2+2,\ \ldots,\ n. \qquad (31.75)$$

При гипотезе H_0 каждое из $n!$ возможных упорядочений значений x имеет одну и ту же вероятность, но независимо от выполнения H_0 $n_1!$ перестановок позиций в первой выборке и $n_2!$ перестановок позиций во второй выборке не влияют на распределение n значений по двум выборкам. Таким образом, имеются $n!/(n_1!n_2!)$ различных распределений по двум выборкам, соответствующих $\binom{n}{n_1}$ способам выбрать элементы первой выборки из n значений.

31.43 Для гипотезы (31.74), в отличие от рассмотренных ранее гипотез (31.9) и (31.49), мы можем рассматривать класс альтернатив гораздо более широкий, чем альтернативы нормальной теории, а именно:

$$H_1\colon\ F_2(x) = F_1(x-\theta) \quad \text{при всех} \quad x. \qquad (31.76)$$

В (31.76) утверждается, что два распределения отличаются только сдвигом. В терминах (31.76) гипотеза (31.74) принимает вид

$$H_0\colon\ \theta = 0. \qquad (31.77)$$

Мы будем называть (31.76) альтернативной гипотезой сдвига. Следует заметить, что, хотя в (31.76) и входит параметр сдвига θ, гипотеза (31.77) является непараметрической согласно нашему определению в **22.3**, поскольку форма распределения $F_1(x)$ не задана.

31.44 Чтобы предложить статистику для проверки H_0, мы вернемся к случаю нормальных альтернатив. Рассмотрим два нормальных распределения, отличающихся только сдвигом. Без потери общности мы предполагаем, что их общая дисперсия σ^2 равна 1, а среднее первого распределения равно нулю. Функция правдоподобия равна, следовательно,

$$L(x|H_1) = (2\pi)^{-\frac{1}{2}n} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n_1} x_{1i}^2 - \frac{1}{2}\sum_{i=1}^{n_2}(x_{2i}-\theta)^2\right\} =$$

$$= (2\pi)^{-\frac{1}{2}n} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \theta\sum_{i=1}^{n_2} x_{2i} - \frac{1}{2}n_2\theta^2\right\}. \qquad (31.78)$$

Из (31.78) мы видим, что при $\theta > 0$ $L(x|H_1)$ достигает максимума, когда $\sum_{i=1}^{n_2} x_{2i}$ принимает наибольшее возможное значение,

и аналогично при $\theta < 0$, когда $\sum_{i=1}^{n_2} x_{2i}$ принимает наименьшее возможное значение. По лемме Неймана — Пирсона из **22.10** наиболее мощная критическая область будет содержать те из $\binom{n}{n_1}$ равновероятных точек в пространстве выборок, которые максимизируют $L(x|H_1)$. Мы, таким образом, приходим к статистике $\sum_{i=1}^{n_2} x_{2i}$ или, что эквивалентно, к среднему второй выборки $\bar{x}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} x_{2i}$. Так как $n_1\bar{x}_1 + n_2\bar{x}_2 = n\bar{x}$ и общее среднее $\bar{x}$ инвариантно при перестановках, то $\bar{x}_2$ определяет также значение $\bar{x}_1$, и будет эквивалентным рассматривать $\bar{x}_1$ или $\bar{x}_1 - \bar{x}_2$. Для двусторонней альтернативы $\theta \neq 0$ мы склонны использовать двусторонний критерий «равных хвостов» над $\bar{x}_2 - \bar{x}_1$ или эквивалентный ему односторонний критерий над $(\bar{x}_1 - \bar{x}_2)^2$ с критической областью, образуемой большими значениями. Именно в такой форме эта статистика критерия была впервые исследована Питмэном (1937a).

Перестановочное распределение w

31.45 Статистика $(\bar{x}_1 - \bar{x}_2)^2$ может принимать значения от нуля до максимума, который достигается, когда каждый член первой выборки равен $\bar{x}_1$ и каждый член второй выборки равен $\bar{x}_2$. Мы тогда получаем для выборочной дисперсии объединенной выборки, скажем s^2, которая инвариантна при перестановках, значение

$$ns^2 = n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 = \frac{n_1 n_2}{n}(\bar{x}_1 - \bar{x}_2)^2.$$

Поэтому мы имеем

$$0 \leqslant (\bar{x}_1 - \bar{x}_2)^2 \leqslant \frac{n^2 s^2}{n_1 n_2}.$$

Следовательно, если мы определим

$$w = \frac{n_1 n_2}{n^2 s^2}(\bar{x}_1 - \bar{x}_2)^2, \tag{31.79}$$

то для всех возможных выборок

$$0 \leqslant w \leqslant 1. \tag{31.80}$$

31.46 Чтобы получить перестановочное распределение статистики w при гипотезе H_0, перепишем ее тождественно как

$$w = \frac{n_1}{n_2 s^2}(\bar{x}_1 - \bar{x})^2, \tag{31.81}$$

В этом выражении при перестановках наблюдений изменяется только $\bar{x}_1$. Точное распределение $\bar{x}_1$ может быть табулировано путем перебора. Но, как ранее отмечалось в **31.19**, этот процесс становится трудоемким при возрастании n. Однако к выражению (31.81) мы можем применить уже полученные результаты, чтобы получить моменты w, поскольку с точностью до множителя это есть квадрат отклонения выборочного среднего от генерального при выборке n_1 элементов из конечной совокупности, содержащей n элементов. В (12.114) и (12.120) мы находим необходимые нам математические ожидания, которые в наших теперешних обозначениях имеют вид

$$\mathbf{M}(\bar{x}_1 - \bar{x})^2 = \frac{(n-n_1)\,s^2}{(n-1)\,n_1} = \frac{n_2 s^2}{n_1(n-1)} \tag{31.82}$$

и

$$\mathbf{M}(\bar{x}_1 - \bar{x})^4 = \frac{n_2}{n_1^3(n-1)(n-2)(n-3)} \times$$
$$\times \{3n(n_1-1)(n_2-1)\,s^4 + [n(n+1) - 6n_1n_2]\,m_4\},$$

где m_4 — выборочный четвертый момент объединенной выборки. Отсюда

$$\mathbf{M}(w) = \frac{1}{n-1}, \tag{31.83}$$

$$\mathbf{M}(w^2) = \frac{1}{n_1n_2(n-1)(n-2)(n-3)} \times$$
$$\times \{3n_1n_2(n-6) + 6n + [n(n+1) - 6n_1n_2]\,g_2\}, \tag{31.84}$$

где g_2 — коэффициент эксцесса $(m_4/s^4) - 3$. Когда какое-либо из чисел n_1 или n_2, а с ним и n, становится большим, (31.84) асимптотически равно

$$\mathbf{M}(w^2) \sim \frac{3}{n^2-1}\left\{1 + O\left(\frac{g_2}{3n_j}\right)\right\}, \tag{31.85}$$

где n_j — объем малой выборки. Если и n_1 и $n_2 \to \infty$, то (31.84) равно

$$\mathbf{M}(w^2) \sim \frac{3}{n^2-1}\left\{1 + O\left(\frac{g_2}{n}\right)\right\}. \tag{31.86}$$

Таким образом, в особенности когда g_2 мало, мы имеем

$$\mathbf{M}(w^2) \sim \frac{3}{n^2-1}. \tag{31.87}$$

31.47 Выражения (31.83) и (31.87) представляют собой первые два момента относительно нуля бета-распределения первого рода:

$$dF = \frac{1}{\mathrm{B}\{1/2,\ (n/2)-1\}}\, w^{-1/2}(1-w)^{(n/2)-2}\,dw, \qquad 0 \leqslant w \leqslant 1, \tag{31.88}$$

и поэтому мы можем ожидать, что оно будет аппроксимировать перестановочное распределение w. И действительно, Питмэн (1937a, b) показал, что третьи моменты также близки и что аппроксимация является очень хорошей.

Рассмотрим теперь обычную статистику Стьюдента t^2 для проверки различия в сдвиге двух нормальных распределений. В наших теперешних обозначениях мы запишем ее как

$$\frac{t^2}{n-2} = \frac{(\bar{x}_1 - \bar{x}_2)^2 \frac{n_1 n_2}{n}}{n_1 s_1^2 + n_2 s_2^2}, \tag{31.89}$$

где s_1^2 и s_2^2 — отдельные выборочные дисперсии. Пользуясь тождеством

$$ns^2 \equiv n_1 s_1^2 + n_2 s_2^2 + \frac{n_1 n_2}{n}(\bar{x}_1 - \bar{x}_2)^2,$$

мы видим из (31.79) и (31.89), что

$$w \equiv \frac{1}{1 + \frac{n-2}{t^2}}. \tag{31.90}$$

Таким образом, мы имеем дело с монотонно возрастающей функцией от t^2. Более того, в точной нормальной теории преобразование (31.90), примененное к распределению Стьюдента с $\nu = n-2$, дает в точности распределение (31.88). (На самом деле мы выполнили, по существу, это же преобразование при сведении функции распределения Стьюдента к неполной бета-функции в **16.11**, с той разницей, что там мы преобразовали к $1-w$ и получили (31.88) с $1-w$ вместо w.)

Мы, таким образом, обнаружили точно так же, как в **31.19**, что приближением к перестановочному распределению при проверке непараметрической гипотезы служит в точности распределение из нормальной теории. В этом частном случае мы можем согласно (31.90) применять w, полагая $(n-2)w/(1-w) = t^2$ с $n-2$ степенями свободы. Для односторонних критериев, обсуждавшихся вначале, мы просто используем подходящий хвост t вместо t^2.

31.48 Широкая применимость t^2 как аппроксимации, даваемой нормальной теорией, в этом примере, несомненно, объясняется действием центральной предельной теоремы, поскольку мы имеем дело с разностью между средними; см. аналогичное обсуждение устойчивости в **31.4**.

Точно так же, как мы отмечали ранее в **31.30**, асимптотическая эквивалентность распределений перестановочного критерия и оптимального критерия нормальной теории влечет за собой, что первый имеет АОЭ, равную 1, против нормальных альтернатив, в данном случае — сдвига.

Свободные от распределения доверительные интервалы для параметра сдвига

31.49 Мы можем теперь использовать статистику критерия w, определяемую в (31.79), чтобы получать свободные от распределения доверительные интервалы для параметра сдвига θ в (31.76). Потому что, каково бы ни было значение θ, (31.76) означает, что n_1 значений x_{1i} $(i = 1, 2, \ldots, n_1)$ и n_2 значений $x_{2i} + \theta$ $(i = 1, 2, \ldots, n_2)$ представляют собой две выборки из одного и того же распределения $F_1(x)$. Следовательно, для этих двух выборок применимо распределение статистики w, отвечающее гипотезе H_0. Обозначим $w(\theta)$ значение w, вычисленное для этих двух выборок и являющееся, очевидно, функцией от θ. Пусть w_α — верхнее критическое значение w для критерия размера α, т. е.

$$\mathbf{P}\{w(\theta) \leqslant w_\alpha\} = 1 - \alpha. \tag{31.91}$$

Вследствие (31.90) (31.91) эквивалентно соотношению

$$\mathbf{P}\{t^2(\theta) \leqslant t_\alpha^2\} = 1 - \alpha, \tag{31.92}$$

где t^2 определено в (31.89). Знаменатель в t^2 есть функция только от выборочных дисперсий отдельных выборок и, следовательно, не зависит от θ. Поэтому, применяя (31.89) к (31.92), имеем

$$\mathbf{P}\{[\bar{x}_1 - (\bar{x}_2 + \theta)]^2 \leqslant k_\alpha^2\} = 1 - \alpha, \tag{31.93}$$

где

$$k_\alpha^2 = \frac{n(n_1 s_1^2 + n_2 s_2^2)}{n_1 n_2 (n-2)} t_\alpha^2. \tag{31.94}$$

Таким образом, мы имеем из (31.93), каково бы ни было истинное значение θ,

$$\mathbf{P}\{(\bar{x}_1 - \bar{x}_2) - k_\alpha \leqslant \theta \leqslant (\bar{x}_1 - \bar{x}_2) + k_\alpha\} = 1 - \alpha, \tag{31.95}$$

и (31.95) является доверительным интервалом для θ.

Если объемы выборок достаточно велики для того, чтобы перестановочное распределение t^2 хорошо приближалось распределением Стьюдента, мы получаем t_α из таблиц последнего; в противном случае нужно пользоваться точным перестановочным распределением w и равенством (31.90). Разумеется, тогда мы можем выбирать для нашего критерия или доверительного интервала только значения α, кратные $\binom{n}{n_1}^{-1}$.

Состоятельность w-критерия

31.50 Используя результат последнего пункта, легко показать, что w — состоятельный критерий для гипотезы (31.77) против альтернативы (31.76) с $\theta \neq 0$ при условии, что $F_1(x)$ имеет конечную дисперсию. На самом деле, как показал Питмэн (1948),

даже если $F_1(x)$ и $F_2(x)$ имеют различные конечные дисперсии и различные средние, критерий w остается состоятельным.

Рассмотрим величину k^2, определенную в (31.94). Если n_1, $n_2 \to \infty$ так, что $n_1/n_2 \to \lambda$, $0 < \lambda < \infty$, то k^2 сходится по вероятности к величине $\dfrac{(\lambda\sigma_1^2 + \sigma_2^2)}{n_1} t_\alpha^2$, имеющей порядок n_1^{-1}, тогда как согласно закону больших чисел $\bar{x}_1 - \bar{x}_2$ сходится к величине $\mathbf{M}(x_1) - \mathbf{M}(x_2)$, равной истинному значению параметра θ, скажем θ_0. Соотношение (31.95) теперь показывает, что для любого α доверительный интервал

$$I = (\bar{x}_1 - \bar{x}_2 - k,\ \bar{x}_1 - \bar{x}_2 + k)$$

есть интервал значений θ, имеющий длину порядка $n_1^{-1/2}$. Если мы выберем $\alpha \leq \varepsilon$, то для любого $\theta_1 \neq \theta_0$

$$\lim_{n_1 \to \infty} \mathbf{P}\{\theta_1 \in I\} < \varepsilon, \tag{31.96}$$

и это всего лишь перевод на язык доверительных интервалов доказываемого утверждения о состоятельности, так как ε может быть сколь угодно близко к нулю. Рано или поздно при возрастании n_1 интервал с вероятностью, стремящейся к 1, не будет содержать θ_1, т. е. критерий для гипотезы $\theta = \theta_0$ будет отвергать $\theta_1 \neq \theta_0$. Это рассуждение делает также очевидным, что (31.77) может быть заменено на H_0: $\theta = \theta_0$, если мы добавим приращение θ_0 к каждому наблюдению второй выборки.

Ранговые критерии для задачи двух выборок

31.51 Точно так же, как в **31.21**, когда мы обсуждали критерии независимости, мы и здесь, в задаче двух выборок, видим, что, если мы хотим иметь возможность табулировать точное перестановочное распределение статистик критерия для любого n, мы должны исключить зависимость статистики критерия от наблюденных значений, которые являются случайными величинами, и это приводит нас к использованию ранговых критериев, которые особенно уместны здесь вследствие их инвариантности относительно монотонных преобразований наблюдаемых величин, оставляющих инвариантной гипотезу (31.74). Снова простейшая процедура состоит в том, чтобы просто заменить наблюдения x_i их рангами, т. е. упорядочить по величине $n_1 + n_2 = n$ наблюдений и заменить значение x_i его рангом X_i. Мы тогда получаем набор из n значений X_i, представляющий собой перестановку первых n натуральных чисел, из которых n_1 относятся к первой выборке и n_2 — ко второй.

Поскольку, как мы указывали в **31.44**, статистика w эквивалентна использованию среднего значения первой выборки $\bar{x}_1$, ранговый критерий, получаемый из w заменой наблюдений их рангами, эквивалентен использованию

$$S = \sum_{i=1}^{n_1} X_i, \tag{31.97}$$

суммы рангов первой выборки, что аналогично r_s в (31.21), так как и то и другое возникает из замены наблюдений рангами.

31.52 Теперь предположим, что мы ищем аналог статистики t, определенной в (31.23), т. е., по существу, статистики Q, определяемой посредством (31.38). Если гипотеза (31.74) справедлива, следует, очевидно, ожидать, что наблюдения из первой и второй выборок будут хорошо перемешаны, так что ранги первой выборки не будут иметь никакой тенденции скапливаться у какого-нибудь или у обоих концов интервала от 1 до n. Определим статистику U, которая считает, сколько раз элемент первой выборки превосходит элемент второй выборки, т. е.

$$U = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} h_{ij}, \tag{31.98}$$

где h_{ij}, как и раньше, определяется формулой (31.36). Статистика U принимает значения от 0 до $n_1 n_2$.

В то время как в случае критериев независимости имеется реальный выбор между r_s и t как статистиками критериев (хотя, как мы видели, они и эквивалентны с точки зрения АОЭ), в двухвыборочной ситуации статистики (31.97) и (31.98) функционально связаны. Действительно,

$$U \equiv S - \tfrac{1}{2} n_1 (n_1 + 1). \tag{31.99}$$

Чтобы доказать (31.99), необходимо только заметить, что

$$U = \sum_{i=1}^{n_1} \left(\sum_{j=1}^{n_2} h_{ij} \right) = \sum_{i=1}^{n_1} (X_i - i).$$

Таким образом, мы можем пользоваться любой из статистик U и S, какая удобнее. Как с теоретической, так и с вычислительной точки зрения статистика U проще.

Статистика S была предложена независимо многими авторами (исторические подробности даются Крускалом и Уоллисом (1952, 1953) и Крускалом (1957)) и называется обычно статистикой критерия Вилкоксона по имени Вилкоксона (1945, 1947); некоторые авторы называют ее статистикой критерия Манна — Уитни по имени авторов, изучавших U несколько позднее (Манн и Уитни, 1947); некоторые называют ее статистикой критерия суммы рангов.

Распределение статистики критерия Вилкоксона

31.53 Мы приступаем к нахождению распределения статистики U, когда справедлива гипотеза (31.74). Мы можем получить ее х. ф. непосредственно из х. ф. статистики Q, даваемой в (31.28). Потому что Q основана на $\frac{1}{2}n(n-1)$ возможных сравнениях между n наблюдениями, тогда как U основана на n_1n_2 сравнениях между первыми n_1 и вторыми n_2 из них, при этом $\frac{1}{2}n_1(n_1-1)$ «внутренних» сравнений для первой выборки и $\frac{1}{2}n_2(n_2-1)$ для второй выборки исключаются. Мы можем записать это соотношение символически как

$$Q_n = Q_{n_1} + Q_{n_2} + U. \tag{31.100}$$

Поскольку при гипотезе H_0 компоненты в правой части (31.100) взаимно независимы, мы имеем соотношение между характеристическими функциями (см. **7.18**)

$$\varphi_n(\theta) = \varphi_{n_1}(\theta)\,\varphi_{n_2}(\theta)\,\Phi_U(\theta),$$

где первые три х. ф. суть х. ф. Q при объемах выборок, указываемых индексами при φ, а $\Phi_U(\theta)$ — х. ф. U. Таким образом,

$$\Phi_U(\theta) = \varphi_n(\theta)/\{\varphi_{n_1}(\theta)\,\varphi_{n_2}(\theta)\}. \tag{31.101}$$

После логарифмирования получаем для п. ф. с.

$$\Psi_U(\theta) = \psi_n(\theta) - \psi_{n_1}(\theta) - \psi_{n_2}(\theta), \tag{31.102}$$

где ψ определено соотношением (31.30). Подставляя его в правую часть (31.102), находим

$$\Psi_U(\theta) = \tfrac{1}{2}n_1n_2i\theta + \sum_{j=1}^{\infty}\frac{B_{2j}(i\theta)^{2j}}{2j\,(2j)!}\left\{\sum_{s=1}^{n}-\sum_{s=1}^{n_1}-\sum_{s=1}^{n_2}\right\}(s^{2j}-1) =$$

$$= \tfrac{1}{2}n_1n_2i\theta + \sum_{j=1}^{\infty}\frac{B_{2j}(i\theta)^{2j}}{2j\,(2j)!}\sum_{s=1}^{n_1}\{(n_2+s)^{2j}-s^{2j}\}. \tag{31.103}$$

Семиинварианты U равны, следовательно,

$$\left.\begin{aligned} \varkappa_1 &= \tfrac{1}{2}n_1n_2, \quad \varkappa_{2j+1} = 0, \quad j \geqslant 1, \\ \varkappa_{2j} &= \frac{B_{2j}}{2j}\sum_{s=1}^{n_1}\{(n_2+s)^{2j}-s^{2j}\}. \end{aligned}\right\} \tag{31.104}$$

В частности, мы находим

$$\varkappa_2 = \frac{1}{12} n_1 n_2 (n+1). \tag{31.105}$$

Соотношение (31.104) показывает, что распределение U симметрично, каковы бы ни были значения n_1 и n_2. Вследствие (31.99) распределение S имеет те же семиинварианты, за исключением того, что ее среднее равно $\frac{1}{2} n_1 (n+1)$.

31.54 Точное распределение S или U может быть легко получено из соотношения между производящими функциями частот, аналогичного (31.101),

$$G_U(\theta) = \binom{n}{n_1}^{-1} G(\theta, n)/\{G(\theta, n_1) G(\theta, n_2)\},$$

где $G_U(\theta)$ — п. ф. ч. для U, а $G(\theta, n)$ — п. ф. ч. для Q, полученная в **31.25**. Подставляя ее значение из (31.27), находим

$$G_U(\theta) = \binom{n}{n_1}^{-1} \frac{\prod_{s=1}^{n} (\theta^s - 1)}{\prod_{s=1}^{n_1} (\theta^s - 1) \prod_{s=1}^{n_2} (\theta^s - 1)} = \binom{n}{n_1}^{-1} \prod_{s=1}^{n_1} \left(\frac{\theta^{n_2+s} - 1}{\theta^s - 1} \right). \tag{31.106}$$

Коэффициент при θ^U во втором сомножителе в правой части (31.106) равен числу способов, которыми можно выбрать n_1 из первых n натуральных чисел так, чтобы их сумма была равна $\frac{1}{2} n_1 (n_1 + 1) + U$. Мы обозначим это число через $f(U, n_1, n_2)$, а накопленную сумму таких чисел — через

$$A(U, n_1, n_2) = \sum_{r=\frac{1}{2} n_1 (n_1+1)}^{U} f(r, n_1, n_2). \tag{31.107}$$

Функция $A(U, n_1, \infty)$ была табулирована Эйлером более двух веков назад. Фикс и Ходжес (1955) дали таблицы, позволяющие вычислять $A(U, n_1, n_2)$ для n_1 (которое без потери общности можно принять $\leqslant n_2$) $\leqslant 12$ и любого n_2. Ранее Уайт (1952) табулировал критические значения S для размеров критерия «равных хвостов» $\alpha = 0{,}05$, 0,01 и 0,001 и $n_1 + n_2 \leqslant 30$. Другие таблицы указаны Крускалом и Уоллисом (1952, 1953) и Фикс и Ходжесом (1955). Фердоорен (1963) и Джейкобсон (1963) приводят расширенные таблицы и поправки к имевшимся ранее таблицам *).

*) В *Таблицах математической статистики* воспроизведена таблица Фердоорена и даны приближенные формулы для вычисления критических значений U в области, не покрываемой таблицей. Таблица меньшего объема приведена в *Сборнике статистических таблиц*, где даются, кроме того, таблицы точного распределения U для $n_2 = 1(1)10$, $n_1 \leqslant n_2$. (*Прим. ред.*)

31.55 Асимптотическая нормальность U при гипотезе H_0 сразу следует из (31.104). Если $n_1, n_2 \to \infty$ так, что $0 < \lim n_1/n_2 = \lambda < \infty$, будем без различия писать N для обозначения n_1, n_2 или n. Тогда $\varkappa_{2j}$ имеет, самое большее, порядок N^{2j+1}, так что (как мы и видели в (31.105)) $\varkappa_2$ имеет порядок N^3. Таким образом,

$$\varkappa_{2j}/(\varkappa_2)^j = O(N^{2j+1-3j}) = O(N^{1-j}),$$

откуда

$$\lim_{N\to\infty} \varkappa_{2j}/(\varkappa_2)^j = 0, \qquad j > 1,$$

и распределение U сходится к нормальному со средним и дисперсией, указанными в (31.104), (31.105). Сходимость к нормальному распределению быстрая, и нормальное приближение эффективно, когда n_1 и n_2 равны примерно 8.

Пользуясь общей теоремой, принадлежащей Гёфдингу (1948а), Леман (1951) показал, что U асимптотически нормально, если две выборки получены из различных непрерывных распределений F_1, F_2, а $\lim n_1/n_2 = \lambda$ ограничен, как и выше.

Леман (1963) получил выражения для свободных от распределения доверительных интервалов, получаемых из статистики критерия Вилкоксона методом пункта **31.49** (см. упражнения 31.24, 31.25).

Состоятельность и несмещенность критерия Вилкоксона

31.56 Из приведенного в **31.50** доказательства состоятельности w-критерия, который переходит в критерий Вилкоксона, когда значения варианты заменяются рангами, следует, что критерий Вилкоксона состоятелен против альтернатив, при которых распределения наблюдений F_1 и F_2 порождают различные средние ранги в своих выборках. Ясно, что это произойдет тогда и только тогда, когда вероятность p того, что наблюдение из F_1 превосходит наблюдение из F_2, отлична от 1/2. Этот результат, полученный Питмэном (1948) и независимо более поздними авторами, может быть доказан непосредственно, как указано в упражнении 31.16.

31.57 Если мы рассмотрим одностороннюю альтернативную гипотезу, что распределение второй выборки «стохастически больше», т. е.

$$H_1\colon F_1(x) > F_2(x) \quad \text{при всех} \quad x, \tag{31.108}$$

то нетрудно доказать, что односторонние критерии Питмэна и Вилкоксона являются несмещенными против (31.108). На самом деле Леман (1951) доказал несмещенность любой подобной критической области для проверки гипотезы (31.74) против (31.108), удовлетворяющей интуитивно желательному условию (C), что

при увеличении величины любого элемента второй выборки выборочная точка остается в критической области. Для любой пары F_1, F_2 определим функцию $h(x)$ равенством

$$F_2(h(x)) = F_1(x), \tag{31.109}$$

так что из (31.108) следует

$$h(x) > x. \tag{31.110}$$

Рассмотрим теперь две выборки, в которых n_2 элементов второй выборки преобразованы из x_i в $h(x_i)$, а первая выборка оставлена без изменений. Мы видим из (31.109), что если преобразованные выборки подчиняются распределениям F_1 и F_2, то до преобразования их распределения были одинаковы. Если теперь область в преобразованном выборочном пространстве имеет вероятность P, то неравенство (31.110) и условие (C), принятое выше, обеспечивают нам, что в непреобразованном выборочном пространстве ее вероятность будет $\alpha \leqslant P$. Поскольку α есть размер критерия над непреобразованными величинами, а P — его мощность против (31.108), то мы видим, что критерий несмещен.

Условию (C), очевидно, удовлетворяют односторонние критерии Питмэна и Вилкоксона.

31.58 Если мы теперь рассмотрим общую двустороннюю альтернативную гипотезу

$$H_1\colon\ F_1(x) \neq F_2(x) \tag{31.111}$$

или даже более ограниченную альтернативу сдвига (31.76) с любым возможным знаком θ, то критерий Вилкоксона, вообще говоря, уже не будет несмещенным. Для альтернатив сдвига Ван дер Ваарт (1950, 1953) показал, что если $n_1 = n_2$ или общая функция плотности симметрична относительно некоторого значения (не обязательно θ), то первая производная функции мощности, если она существует, равна нулю при $\theta = 0$, но даже в этом случае критерий может не быть несмещенным.

АОЭ критерия Вилкоксона

31.59 Мы теперь ограничимся ситуацией сдвига (31.76) и найдем АОЭ критерия Вилкоксона по сравнению с t-критерием Стьюдента (который является оптимальным критерием сдвига, если F_1 — нормальное распределение), когда F_1 — произвольная непрерывная ф. р. с конечной дисперсией.

Критерий t Стьюдента, определенный в (31.89), как известно, асимптотически эквивалентен использованию статистики $(\bar{x}_1 - \bar{x}_2)$, которая, какова бы ни была ф. р. F_1, асимптотически нормальна со средним $-\theta$ и дисперсией $\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right) = \frac{n\sigma^2}{n_1 n_2}$.

Таким образом, мы имеем

$$\frac{\left\{\left[\frac{\partial}{\partial\theta}\mathbf{M}(\bar{x}_1-\bar{x}_2)\right]_{\theta=0}\right\}^2}{\mathbf{D}\{(\bar{x}_1-\bar{x}_2)\mid\theta=0\}}=\frac{n_1n_2}{n\sigma^2}. \tag{31.112}$$

Нам теперь нужно оценить аналог (31.112) для статистики Вилкоксона U. Из определения U в (31.98) имеем

$$\mathbf{M}(U)=n_1n_2\mathbf{M}(h_{ij})=n_1n_2p,$$

где p — вероятность того, что наблюдение x_1 из распределения $F_1(x)$ превосходит наблюдение x_2 из распределения $F_1(x-\theta)$. Пользуясь формулой (11.68) для ф. р. суммы двух случайных величин (с заменой $-y$ на y в аргументе F_1 и перестановкой индексов 1, 2, чтобы получить ф. р. x_2-x_1), находим

$$p=H(0)=\int_{-\infty}^{\infty}F_1(0+x-\theta)f_1(x)\,dx,$$

откуда

$$\frac{\partial p}{\partial\theta}=-\int_{-\infty}^{\infty}f_1(x-\theta)f_1(x)\,dx$$

и

$$\left[\frac{\partial\mathbf{M}(U)}{\partial\theta}\right]_{\theta=0}=n_1n_2\int_{-\infty}^{\infty}\{f_1(x)\}^2\,dx. \tag{31.113}$$

Соотношения (31.113) и (31.105) дают

$$\frac{\left\{\left[\frac{\partial\mathbf{M}(U)}{\partial\theta}\right]_{\theta=0}\right\}^2}{\mathbf{D}(U\mid\theta=0)}=\frac{12n_1n_2}{n+1}\left[\int_{-\infty}^{\infty}\{f_1(x)\}^2\,dx\right]^2. \tag{31.114}$$

Пользуясь (25.27), (31.112) и (31.114), имеем для АОЭ критерия U Вилкоксона по сравнению с критерием t Стьюдента

$$A_{U,t}=12\sigma^2\left[\int_{-\infty}^{\infty}\{f_1(x)\}^2\,dx\right]^2. \tag{31.115}$$

Этот результат принадлежит Питмэну (1948).

31.60 Чтобы вычислить (31.115), нам требуется лишь значение интеграла

$$\int_{-\infty}^{\infty}\{f_1(x)\}^2\,dx=\mathbf{M}\{f_1(x)\}. \tag{31.116}$$

Значение (31.116) легко находится для конкретных распределений.

Когда F_1 нормальна, мы имеем

$$\mathbf{M}\{f_1(x)\} = (2\pi\sigma^2)^{-1/2}\,\mathbf{M}\left\{\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]\right\} =$$

$$= (2\pi\sigma^2)^{-1/2}\left[\mathbf{M}\left\{\exp\left[it\left(\frac{x-\mu}{\sigma}\right)^2\right]\right\}\right]_{it=-1/2} =$$

$$= (2\pi\sigma^2)^{-1/2}\left[(1-2it)^{-1/2}\right]_{it=-1/2} = (4\pi\sigma^2)^{-1/2}.$$

Таким образом, из (31.115) в нормальном случае имеем

$$A_{U,t} = \frac{3}{\pi} \approx 0{,}95. \tag{31.117}$$

Этот результат не зависит от масштаба, как это и будет всегда, поскольку $\sigma\mathbf{M}\{f_1(x)\}$ не зависит от масштаба. Поэтому мы всегда можем изменить масштаб в (31.115), если это удобно.

31.61 Легко видеть, что выражение (31.115) может принимать бесконечные значения (см. упражнение 31.18). Теперь нас интересует, существует ли ненулевой минимум, ниже которого оно не может попасть. Мы будем минимизировать $\mathbf{M}\{f(x)\}$ при фиксированном σ^2, которое можем принять равным единице.

Не теряя общности, примем также $\mathbf{M}(x) = 0$. Таким образом, нам требуется минимизировать $\int_{-\infty}^{\infty} f^2(x)\,dx$ при условиях $\int_{-\infty}^{\infty} x^2 f(x)\,dx = 1$, $\int_{-\infty}^{\infty} f(x)\,dx = 1$. Пользуясь неопределенными множителями Лагранжа λ, μ, мы получаем, что это эквивалентно минимизации интеграла

$$\int_{-\infty}^{\infty} \{f^2(x) - 2\lambda(\mu^2 - x^2) f(x)\}\,dx. \tag{31.118}$$

Поскольку $f(x)$ неотрицательна, (31.118) достигает минимума при таком выборе $f(x)$, когда

$$f(x) = \begin{cases} \lambda(\mu^2 - x^2), & x^2 \leqslant \mu^2, \\ 0, & x^2 > \mu^2 \end{cases} \tag{31.119}$$

(параболическое распределение). Определяя λ и μ из условий

$$\int f(x)\,dx = \int x^2 f(x)\,dx = 1,$$

находим

$$\mu^2 = 5, \quad \lambda = 3/(20\sqrt{5}), \tag{31.120}$$

откуда

$$\int f^2(x)\,dx = \frac{16}{15}\lambda^2\mu^5 = \frac{3}{5\sqrt{5}}. \tag{31.121}$$

Таким образом, из (31.121) и (31.115) АОЭ для распределения (31.119), (31.120) равна

$$\inf A_{U,t} = 108/125 = 0{,}864. \qquad (31.122)$$

31.62 Высокое значение (31.122) для минимума АОЭ критерия U Вилкоксона по сравнению с критерием t Стьюдента, которое было впервые получено Ходжесом и Леманом (1956), является хорошим свидетельством в пользу критерия U. В практических терминах оно означает, что при больших выборках мы не можем потерять более 13,6% эффективности при использовании критерия Вилкоксона вместо критерия Стьюдента для альтернатив сдвига; с другой стороны, мы можем получить большой выигрыш (см. упражнение 31.18, из которого можно вывести, что для гамма-распределения с параметром $p = 1$ $A_{U,t} = 3$). Если распределение в действительности нормально, потеря эффективности составляет только около 5% согласно (31.117). Ван дер Ваарт (1950) показал, что если в нормальном случае рассматривать отношения производных функций мощности критериев Вилкоксона и Стьюдента, то их значения при очень небольших объемах выборок мало отличаются от асимптотических значений. Сандрум (1953) вычислил приближенные функции мощности в нормальном и равномерном случаях при $n_1 = n_2 = 10$, которые подтверждают, в частности, в нормальном случае малость потери мощности, вызванной использованием критерия Вилкоксона вместо критерия Стьюдента. Диксон (1954) и Ходжес и Леман (1956) приводят некоторые результаты для малых выборок в нормальном случае, которые подтверждают это. Уиттинг (1960), используя разложение Эджворта до порядка n^{-2}, показал, что значение (31.117) для АОЭ в нормальном случае сохраняется с большой точностью для объемов выборок от 4 до 40.

Чанда (1963) нашел высокую АОЭ даже для дискретных генеральных совокупностей. Нётер (1967) показал, что критерий и основанные на нем доверительные интервалы консервативны для дискретных генеральных совокупностей. Хейнэм и Говинд-араджулу (1966) приводят точные вычисления мощности в экспоненциальном и равномерном случаях.

Критерий, имеющий равномерно лучшую АОЭ по сравнению с критерием Стьюдента

31.63 Хотя, как мы видели в **31.59—62**, критерий Вилкоксона имеет очень хорошие качества по сравнению с критерием Стьюдента, можно найти еще лучший. Возвращаясь к нашему обсуждению в **31.39**, мы можем, очевидно, получить АОЭ, равную 1, против *нормальных* альтернатив сдвига, используя статистику w

(или, что эквивалентно, среднее первой выборки $\bar{x}_1$), в которой наблюдения заменены математическими ожиданиями порядковых статистик в нормальных выборках, которые мы обозначим $E(s, n)$, как и в **31.39**. Статистика критерия тогда эквивалентна

$$c_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} E(X_i, n), \tag{31.123}$$

где X_i — ранг i-го наблюдения первой выборки среди n наблюдений. (31.123) обычно называют статистикой нормальных меток или статистикой критерия c_1. Если мы положим

$$z_s = \begin{cases} 1, & \text{если } s\text{-е наблюдение принадлежит первой выборке,} \\ 0 & \text{в противном случае,} \end{cases}$$

то мы можем переписать (31.123) как

$$c_1 = \frac{1}{n_1} \sum_{s=1}^{n} E(s, n) z_s. \tag{31.124}$$

Гёфдинг (1952) впервые показал, что c_1-критерий имеет АОЭ 1 при нормальных альтернативах. Терри (1952) дал точное распределение c_1 для $n_1 + n_2 \leqslant 10$, а Клоц (1964) получил критические значения для $n_1 + n_2 \leqslant 20$ при гипотезе (31.74). Асимптотическая нормальность c_1 (и широкого класса аналогичных статистик) для любых распределений F_1, F_2, когда $\lim n_1/n_2$ отличен от 0 и ∞, была доказана Черновым и Сэвиджем (1958). Ясно, что

$$\mathsf{M}(c_1 \mid H_0) = \frac{1}{n_1} \sum_{s=1}^{n} E(s, n)\, \mathsf{M}(z_s) = \frac{1}{n_1} \cdot \frac{n_1}{n} \sum_{s=1}^{n} E(s, n) = 0;$$

$\mathsf{D}(c_1 \mid H_0)$ указана в (31.133) ниже.

31.64 Для целей вычисления АОЭ c_1 более удобно иное ее определение. Определим функцию распределения объединения исходных генеральных совокупностей

$$H(x) = \frac{n_1}{n} F_1(x) + \frac{n_2}{n} F_2(x) \tag{31.125}$$

и ее выборочный аналог

$$H_n(x) = \frac{n_1}{n} S_{n_1}(x) + \frac{n_2}{n} S_{n_2}(x), \tag{31.126}$$

где S_{n_1} и S_{n_2} — выборочные функции распределения, опреде-

ленные в (30.97). Если мы введем также функцию $J_n(x)$ с помощью равенства

$$E(s, n) = J_n\left(\frac{s}{n}\right), \tag{31.127}$$

то можно определить c_1 в интегральной форме:

$$c_1 = \int_{-\infty}^{\infty} J_n\{H_n(x)\}\, dS_{n_1}(x). \tag{31.128}$$

При $n \to \infty$ мы имеем из (31.125), (31.126) $H_n(x) \to H(x)$, в то время как из (31.127) при слабых ограничениях (см. (31.73)) следует $J_n(x) \to \Phi^{-1}(x) = J(x)$, где Φ — функция стандартного нормального распределения. Таким образом, при $n \to \infty$ имеем из (31.128) при условиях регулярности

$$\mathbf{M}(c_1) = \int_{-\infty}^{\infty} J\{H(x)\}\, dF_1(x). \tag{31.129}$$

Ван дер Варден (1952, 1953) предложил двухвыборочный критерий, в котором вместо $E(s, n)$ используются числа $J\left(\frac{s}{n+1}\right)$, асимптотически эквивалентный критерию c_1 (ср. (31.129)).

31.65 Если мы теперь рассмотрим альтернативы сдвига (31.76) и продифференцируем (31.129) по параметру сдвига θ, мы получим

$$\frac{\partial \mathbf{M}(c_1)}{\partial \theta} = \int_{-\infty}^{\infty} J'\{H(x)\}\left\{\frac{d}{d\theta} H(x)\right\} dF_1(x). \tag{31.130}$$

Так как

$$H(x) = \frac{n_1}{n} F_1(x) + \frac{n_2}{n} F_1(x - \theta),$$

мы имеем

$$\frac{d}{d\theta} H(x) = -\frac{n_2}{n} f_1(x - \theta).$$

Подставляя это в (31.130), находим

$$\frac{\partial \mathbf{M}(c_1)}{\partial \theta} = -\frac{n_2}{n_1} \int_{-\infty}^{\infty} J'\{H(x)\} f_1(x - \theta)\, dF_1(x). \tag{31.131}$$

Поскольку $H(x) = F_1(x)$ при $\theta = 0$, из (31.131) следует

$$\left[\frac{\partial \mathbf{M}(c_1)}{\partial \theta}\right]_{\theta=0} = -\frac{n_2}{n} \int_{-\infty}^{\infty} J'\{F_1(x)\}\{f_1(x)\}^2\, dx. \tag{31.132}$$

При гипотезе H_0 дисперсия c_1 в силу (31.124) равна

$$\mathsf{D}(c_1 \mid H_0) = \frac{1}{n_1^2}\mathsf{D}\left\{\sum_{s=1}^{n} E(s,\ n)\, z_s\right\}.$$

Здесь случайными величинами являются только z_s; это (0—1)-величины, принимающие значение 1 с вероятностью $\frac{n_1}{n}$, но они не независимы и каждая пара имеет корреляцию $-\frac{1}{n-1}$ вследствие симметрии. Таким образом,

$$\mathsf{D}(c_1 \mid H_0) =$$

$$= \frac{1}{n_1^2}\left[\sum_{s=1}^{n}\{E(s,\ n)\}^2\,\mathsf{D}z_s + \sum\sum_{s\neq r} E(s,\ n)\,E(r,\ n)\,\mathrm{cov}(z_s,\ z_r)\right] =$$

$$= \frac{\mathsf{D}z}{n_1^2}\left\{\sum_{s=1}^{n}\{E(s,\ n)\}^2 - \frac{1}{n-1}\left[\left(\sum_{s=1}^{n} E(s,\ n)\right)^2 - \sum_{s=1}^{n}\{E(s,\ n)\}^2\right]\right\},$$

и поскольку $\sum\limits_{s=1}^{n} E(s,\ n) = 0$ вследствие симметрии нормального распределения, а $\mathsf{D}z = \frac{n_1}{n}\left(1 - \frac{n_1}{n}\right)$, то

$$\mathsf{D}(c_1 \mid H_0) = \frac{n_2}{n(n-1)\,n_1}\sum_{s=1}^{n}\{E(s,\ n)\}^2. \qquad (31.133)$$

Это точное выражение. Когда $n \to \infty$, вследствие (31.73)

$$\frac{1}{n}\sum_{s=1}^{n}\{E(s,\ n)\}^2 \to \int x^2\, d\Phi(x) = 1$$

и

$$\mathsf{D}(c_1 \mid H_0) \to \frac{n_2}{(n-1)\,n_1}. \qquad (31.134)$$

Таким образом, из (31.132) и (31.134) имеем

$$\frac{\left\{\left[\frac{\partial}{\partial\theta}\mathsf{M}(c_1)\right]_{\theta=0}\right\}^2}{\mathsf{D}(c_1 \mid \theta = 0)} \to \frac{n_1 n_2}{n}\left[\int_{-\infty}^{\infty} J'\{F_1(x)\}\{f_1(x)\}^2\, dx\right]^2. \qquad (31.135)$$

Пользуясь (25.27), получаем из (31.135) и (31.112), что АОЭ критерия c_1 по сравнению с критерием t Стьюдента равна

$$A_{c_1,\,t} = \left[\int_{-\infty}^{\infty} J'\{F_1(x)\}\{f_1(x)\}^2\, dx\right]^2. \qquad (31.136)$$

В (31.136) мы положили $\sigma^2 = 1$, нормировав $F_1(x)$. Теперь мы будем искать минимальное значение (31.136).

31.66 Вспомним теперь, что $J(x)$ в (31.136) определена как $\Phi^{-1}(x)$, обратная функция к стандартной нормальной ф. р. Поэтому

$$\Phi\{J(x)\} = x. \tag{31.137}$$

Дифференцируя обе части и обозначая $\varphi(x) = \Phi'(x)$ функцию плотности нормального распределения, имеем

$$\frac{d}{dx}[\Phi\{J(x)\}] = \varphi\{J(x)\}J'(x) = 1,$$

так что

$$J'(x) = \frac{1}{\varphi\{J(x)\}}. \tag{31.138}$$

Из (31.137)

$$\Phi\{J[F_1(x)]\} = F_1(x),$$

так что мы имеем уравнение в дифференциалах

$$dF_1(x) = d\Phi\{J[F_1(x)]\} = \varphi\{J[F_1(x)]\}\,dJ[F_1(x)]. \tag{31.139}$$

Равенство (31.138) дает также

$$J'\{F_1(x)\} = \frac{1}{\varphi\{J[F_1(x)]\}},$$

и интеграл в квадратных скобках в (31.136), следовательно, равен

$$I = \int_{-\infty}^{\infty} \frac{1}{\varphi\{J\}} \frac{[\varphi\{J\}\,dJ]^2}{dx} = \int_{-\infty}^{\infty} \varphi\{J\} \Big/ \left(\frac{dx}{dJ}\right) dJ. \tag{31.140}$$

Чтобы минимизировать (31.136), нам нужно минимизировать (31.140) при условиях нормировки

$$\int x\,dF_1 = \int x\varphi\{J\}\,dJ = 0, \quad \int x^2\,dF_1 = \int x^2\varphi\{J\}\,dJ = 1. \tag{31.141}$$

Минимизация производится по F_1, не выписанному явно аргументу функции J в (31.140). Это эквивалентно минимизации по x как функции от J. Мы, следовательно, ищем монотонно неубывающую функцию $x(J)$, которая минимизирует (31.140) при условиях (31.141).

31.67 Поскольку φ — ф. п. нормированного нормального распределения, то очевидно, что ограничения (31.141) удовлетворяются, когда

$$x(J) = J, \tag{31.142}$$

и в этом случае (31.140) равно 1. Согласно (31.139) это будет иметь место, когда $dF_1(x) = d\Phi(x)$, т. е. когда F_1 нормальна.

Таким образом, мы проверили наше утверждение из **31.63** о том, что в нормальном случае критерий c_1 имеет АОЭ 1.

Чернов и Сэвидж (1958) показали, что (31.142) является в действительности единственным решением нашей задачи минимизации, т. е. что $I > 1$ для любой ф. р. F_1, отличной от нормальной, так что минимальная АОЭ критерия c_1 по сравнению с критерием t равна 1. Мы приводим только эвристические рассуждения, ведущие к этому результату. Будем использовать представление $x(J) = J + p(J)$.

Если выполняется (31.141), мы имеем

$$1 = \mathsf{D}J = \mathsf{D}\{x(J)\} = \mathsf{D}J + \mathsf{D}\{p(J)\} + 2\,\mathrm{cov}\{J,\ p(J)\},$$

и, следовательно,

$$\mathrm{cov}\{J,\ p(J)\} \leqslant 0, \tag{31.143}$$

причем равенство в (31.143) достигается тогда и только тогда, когда $p(J)$ постоянна. В силу (31.141) $\mathsf{M}\{p(J)\} = 0$, так что (31.143) обращается в равенство, только когда $p(J) \equiv 0$. Этот случай, рассмотренный в (31.142), мы оставим в стороне. Мы хотим минимизировать выражение (31.140), равное $\mathsf{M}\left\{\frac{1}{x'(J)}\right\}$. Предположим, что $x(J)$ строго возрастает, так что $x'(J) > 0$. Тогда, пользуясь упражнением 9.13, имеем, исключая вырожденный случай равенства,

$$\mathsf{M}\left\{\frac{1}{x'(J)}\right\} > \frac{1}{\mathsf{M}\{x'(J)\}} = \frac{1}{1 + \mathsf{M}\{p'(J)\}}.$$

Но из (31.143) следует при определенных условиях, что $\mathsf{M}\{p'(J)\} < 0$, так как если J и $p(J)$ имеют отрицательную ковариацию, то «средний наклон» $p(J)$ должен быть отрицательным. Поэтому, если $p(J)$ не равно тождественно нулю, то

$$\mathsf{M}\left\{\frac{1}{x'(J)}\right\} > 1,$$

и, следовательно, $x(J) = J$ дает единственный минимум АОЭ.

Микульский (1963) показал при некоторых условиях регулярности, что никакое другое распределение, кроме нормального, не обладает этим замечательным свойством, что эффективность наилучшего рангового критерия по сравнению с наилучшим критерием для сдвига превышает единицу всегда, когда распределение наблюдений отличается от того, которое предполагалось при построении критериев.

Д. Кокс (1964) рассматривал аналогичные применения экспоненциальных меток, полученных в упражнении 19.11.

31.68 Из результата пунктов **31.63—67** следуют важные выводы. Критерий c_1 свободен от распределения (т. е. полностью устойчив), и по сравнению с критерием стандартной нормальной теории, основанным на статистике t Стьюдента, устойчивость которого можно считать лишь достаточно хорошей, его

наименьшая АОЭ равна 1. Трудно поэтому найти подходящую ситуацию, когда было бы оправдано общепринятое использование критерия Стьюдента для проверки сдвига между двумя выборками, если объемы выборок достаточно велики. Критерий t не имеет заметного преимущества (даже в нормальном случае, для которого он оптимален) при объемах выборки 4 или 5 с α, близким 0,05 (см. Клоц (1964)).

Применение критерия c_1 требует очень небольших затрат труда. Для этого нужно вычислить

$$\frac{c_1 - \mathbf{M}(c_1 \mid H_0)}{\{\mathbf{D}(c_1 \mid H_0)\}^{1/2}} = \left\{\frac{n(n-1)}{n_1 n_2}\right\}^{1/2} \frac{\sum_{i=1}^{n_1} E\{X_i, n\}}{\left[\sum_{s=1}^{n} \{E(s, n)\}^2\right]^{1/2}}$$

и обратиться к таблице стандартного нормального распределения. В таблицах Фишера и Иэйтса даются значения $\sum_{s=1}^{n} \{E(s, n)\}^2$ для $n = 1(1)50$ с четырьмя десятичными знаками и значения $E(s, n)$ с двумя знаками. Хартер (1961a) приводит все $E(s, n)$ с пятью десятичными знаками для $n = 2(1)100(25)250(50)400$. При $n = 50$ $\left[\frac{1}{n}\sum_{s=1}^{n} \{E(s, n)\}^2\right]^{1/2} = 0{,}97$ и в дальнейшем стремится к 1, как видно из формулы, следующей за (31.133), так что этот множитель может быть опущен, и тогда нормированная статистика критерия сводится к

$$\left(\frac{n-1}{n_1 n_2}\right)^{1/2} \sum_{i=1}^{n_1} E(X_i, n).$$

Ходжес и Леман (1963) показали, что из критериев Вилкоксона и нормальных меток могут быть получены устойчивые оценки, имеющие такие же эффективности, как АОЭ критериев (см. также Хёйланд (1965)). Рамачандрамурти (1966) дает вид этих оценок, которые устойчивы к различию в масштабе и, следовательно, применимы для общей задачи о двух средних.

Прэтт (1964) изучал влияние неизвестного различия в масштабе на размер критериев Вилкоксона, нормальных меток, Стьюдента и других критериев сдвига. Асимптотически критерий Стьюдента заметно более устойчив лишь при n_1/n_2, очень близком к 1, и даже тогда его преимущество мало, если параметры масштаба отличаются не более чем в 2 раза.

31.69 Были предложены и другие критерии для задачи двух выборок, хотя теперь их несколько заслонили критерии Вилкоксона и c_1. Критерий Вальда и Волфовица (1940) (см. упражнение 30.8), основанный на числе серий, имеет то преимущество, что он состоятелен против общих альтернатив (31.111), если $\lim n_1/n_2$ отличен от 0 или ∞. Смирнов (1939b) предложил критерий, основанный на максимуме абсолютной разности d между двумя эмпирическими функциями распределения, и показал, что $d \Big/ \left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{1/2}$

имеет то же предельное распределение (30.132), что и $D_n n^{1/2}$. Сходимость эмпирических ф. р. к теоретическим обеспечивает состоятельность критерия против альтернатив (31.111). Нижняя граница для его мощности может быть получена точно так же, как для D_n в **30.60**, но критерий может быть смещенным (см. упражнение 30.16 для критерия D_n).

Леман (1951) предложил критерий, который, как он показал, всегда является несмещенным.

Хотя о мощности этих критериев известно довольно мало, ясно, что они менее эффективны против нормальных альтернатив сдвига. Действительно, Муд (1954) показал, что критерий Вальда — Волфовица имеет АОЭ, равную нулю, против нормальных альтернатив различия в сдвиге или масштабе.

Альтернатива различия в масштабе

$$F_2(x) = F_1\left(\frac{x}{\theta}\right), \qquad \theta > 0, \tag{31.144}$$

при которой мы проверяем H_0: $\theta = 1$, эквивалентна альтернативе сдвига для логарифмов величин, если они неотрицательны. Но в общем случае нужно искать новые критерии против альтернатив (31.144). Муд (1954) предложил критерий $W = \sum_{i=1}^{n_1} \left\{ X_i - \frac{1}{2}(n+1) \right\}^2$, свободный от распределения, и показал, что в нормальном случае его АОЭ по сравнению с оптимальным критерием отношения дисперсий равна $15/(2\pi^2) \approx 0{,}76$. Для других распределений его АОЭ принимает значения от 0 до ∞. Если (31.144) заменить более общей альтернативой

$$F_2(x - \mu) = F_1\{(x - \nu)/\theta\}, \tag{31.145}$$

где μ и ν — неизвестные (мешающие) параметры сдвига, критерий Муда остается *асимптотически* свободным от распределения, если μ и ν предварительно оцениваются выборочными медианами (см. Крауз (1964)).

Мы можем получить АОЭ, равную 1, против нормальных альтернатив, применяя критерий отношения дисперсий к нормальным меткам $E(X_i, n)$ (см. Кейпон (1961), где использовалась асимптотически эквивалентная статистика, определяемая формулой (31.123) с $E(X_i, n)$, замененными их квадратами). Клоц (1962) использовал квадрат преобразования посредством обратной нормальной ф. р., что дает также асимптотически эквивалентный критерий (ср. критерий Ван дер Вардена в **31.64**), и табулировал критические значения этой статистики для $8 \leqslant n_1 + n_2 \leqslant 20$. Рагхавачари (1965a, b) подтвердил, что АОЭ критерия Клоца изменяется от 0 до ∞ в зависимости от F в (31.144), и показал, что его асимптотические свойства против (31.144) сохраняются и против (31.145), когда μ, ν подходящим образом состоятельно оцениваются по выборкам, при условии, что F симметрична и достаточно регулярна.

Сиджел и Тьюки (1960) предложили критерий для различия в масштабе, имеющий то же самое H_0-распределение, что и критерий Вилкоксона (поскольку он использует совокупность рангов; они приписываются наблюдениям в соответствии с их расстоянием от концов упорядоченного ряда), но его АОЭ равна лишь $6/\pi^2 = 0{,}61$ в нормальном случае, как показал Клоц (1962).

Ван Эден (1964) дает условия состоятельности этих и других масштабных критериев. Тамура (1963) показал, что против нормальных альтернатив масштаба АОЭ критерия Муда может быть повышена до 0,92 применением пятых степеней вместо квадратов в его определении и что для критерия,

эквивалентного критерию Сиджела — Тьюки, АОЭ может быть повышена до 0,96 использованием очень малых степеней рангов вместо самих рангов.

В работе Мозеса (1963) содержится общее обсуждение критериев различия в масштабе.

k-выборочные критерии

31.70 Обобщение двухвыборочных критериев на задачу k выборок (задача (2) в **31.13**) не вызывает затруднений. На основе независимых выборок, состоящих из n_p наблюдений ($p = 1, 2, \ldots, k$), где $\sum_{p=1}^{k} n_p = n$, проверяется гипотеза, что $k \geqslant 2$ непрерывных функций распределения одинаковы, т. е. гипотеза

$$H_0\colon\ F_1(x) = F_2(x) = \ldots = F_k(x) \quad \text{при всех} \quad x. \tag{31.146}$$

В параметрической теории все F_p предполагаются нормальными с одной и той же неизвестной дисперсией σ^2 и различными средними θ_p, так что

$$H_1\colon\ F_p(x) = F(x - \theta_p), \qquad p = 1, 2, \ldots, k, \tag{31.147}$$

где не все θ_p равны между собой. Тогда (31.146) принимает вид

$$H_0\colon\ \theta_p = 0 \quad \text{при всех} \quad p. \tag{31.148}$$

Критерий отно́шения правдоподобия для (31.148) в нормальном случае основан на статистике

$$F = \frac{1}{k-1}\sum_{p=1}^{k} n_p(\bar{x}_p - \bar{x})^2 \Big/ \frac{1}{n-k}\sum_{p=1}^{k}\sum_{j=1}^{n_p}(x_{pj} - \bar{x}_p)^2$$

(где $\bar{x}_p$ есть среднее p-й выборки, а $\bar{x}$ — среднее всей выборки в целом), которая имеет распределение отношения дисперсий (F) с $(k-1, n-k)$ степенями свободы. Это следует непосредственно из общей теории отношения правдоподобия главы 24 и представляет собой простейший случай дисперсионного анализа, который мы будем рассматривать в томе 3. При $(n-k) \to \infty$ получаем (см. **16.22** (7)), что

$$S = \sum_{p=1}^{k} n_p(\bar{x}_p - \bar{x})^2/\sigma^2 \tag{31.149}$$

имеет асимптотически распределение χ^2 с $k-1$ степенями свободы.

Как показал Гейен (1949—1951), этот критерий чрезвычайно устойчив к отклонениям от нормальности, и общее обсуждение устойчивости процедур дисперсионного анализа будет проводиться в томе 3. Здесь мы рассмотрим только свободные от распределения заменители для критерия нормальной теории.

31.71 Поскольку (31.147) очевидно, является обобщением альтернативы сдвига (31.76) в случае двух выборок, естественно искать обобщения двухвыборочных критериев на k выборок. Мы рассмотрим два различных подхода к этой задаче. Во-первых, предположим, что мы просто заменили наблюдения x их рангами X. Статистика (31.149) тогда принимает вид

$$S = \sum_{p=1}^{k} n_p \left\{ \bar{X}_p - \tfrac{1}{2}(n+1) \right\}^2 \Big/ \left\{ \tfrac{1}{12}(n^2-1) \right\}, \qquad (31.150)$$

приводясь к статистике Вилкоксона, когда $k = 2$. Крускал и Уоллис (1952, 1953) предложили статистику $H = (n-1)S/n$, большие значения которой составляют критическую область. Они показали, что она имеет асимптотически распределение χ^2_{k-1}, как и в параметрическом случае. При $k = 3$, $n_p \leqslant 5$ они табулировали ее точное распределение в окрестности критических значений для $\alpha = 0{,}10,\ 0{,}05,\ 0{,}01$ *). Крускал (1952) показал, что H-критерий состоятелен против любой альтернативной гипотезы, при которой наблюдение из какого-нибудь из распределений имеет вероятность $\neq 1/2$ превзойти случайное наблюдение из k распределений, взятых вместе. Это обобщает условие состоятельности критерия Вилкоксона в **31.56**.

Пури (1964) рассматривал статистику, полученную заменой наблюдений *любым* множеством условных чисел; он дал условия ее асимптотической нормальности и получил ее АОЭ, которая не зависит от k. В отношении критерия H это означает, что его АОЭ дается тем же выражением (31.115), что и для критерия Вилкоксона (см. Эндрюс (1954)), а анализ из **31.60—61** применим без изменений к критерию H. Если вместо рангов используются нормальные метки, мы получаем обобщение критерия c_1 из (31.124), АОЭ дается формулой (31.136) и согласно **31.66—67** она не меньше 1, как и раньше.

31.72 В задаче двух выборок мы видели в **31.52**, что при выводе статистики критерия было безразлично, заменять ли наблюдения рангами или считать число инверсий между выборками. В случае k выборок это не одно и то же. Мы теперь рассмотрим подход, использующий инверсии, и увидим, что он приводит к статистике, отличной от H из **31.71**.

Предположим, что статистика U из (31.98) вычисляется для каждой пары выборок, причем всего таких пар $\frac{1}{2}k(k-1)$; обо-

*) Эта таблица воспроизведена в *Сборнике статистических таблиц*. (*Прим. ред.*)

значим через U_{pq} значение, полученное для p-й и q-й выборок $(p, q = 1, 2, \ldots, k;\ p \neq q)$, и положим

$$U = \sum_{p=1}^{k} \sum_{q=p+1}^{k} U_{pq}. \tag{31.151}$$

Мы теперь легко можем обобщить теорию из **31.53**. Соотношение (31.100) заменяется на $Q_n = \sum_{p=1}^{k} Q_{n_p} + U$, что приводит к соотношению между х. ф.

$$\Phi_U(\theta) = \varphi_n(\theta) \Big/ \prod_{p=1}^{k} \varphi_{n_p}(\theta), \tag{31.152}$$

являющемуся аналогом (31.101). Для п. ф. с. U мы находим, соответственно с (31.103),

$$\Psi_U(\theta) = \frac{1}{4} i\theta \left(n^2 - \sum_{p=1}^{k} n_p^2 \right) + \sum_{j=1}^{\infty} \frac{B_{2j} (i\theta)^{2j}}{2j\,(2j)!} \left\{ \sum_{s=1}^{n} - \sum_{p=1}^{k} \sum_{s=1}^{n_p} \right\} s^{2j}. \tag{31.153}$$

Семиинварианты U равны, следовательно,

$$\left.\begin{aligned} \varkappa_1 &= \frac{1}{4} \left(n^2 - \sum_{p=1}^{k} n_p^2 \right), \quad \varkappa_{2j+1} = 0, \quad j > 0, \\ \varkappa_{2j} &= \frac{B_{2j}}{2j} \left\{ \sum_{s=1}^{n} - \sum_{p=1}^{k} \sum_{s=1}^{n_p} \right\} s^{2j}. \end{aligned}\right\} \tag{31.154}$$

В частности,

$$\varkappa_2 = \frac{1}{72} \left\{ n^2 (2n + 3) - \sum_{p=1}^{k} n_p^2 (2n_p + 3) \right\}. \tag{31.155}$$

31.73 Предельное распределение U выводится так же, как в **31.55**. Если $n_p \to \infty$ так, что n/n_p остаются ограниченными при всех p, то, обозначая через N любое из n_p или n, мы видим, что $\varkappa_{2j}$ имеет самое большее порядок N^{2j+1}, а $\varkappa_2$ — порядок N^3. Поэтому $\varkappa_{2j}/\varkappa_2^j$ имеет порядок самое большее N^{1-j} и стремится к нулю при всех $j > 1$, так что U асимптотически нормально со средним и дисперсией, даваемыми в (31.154), (31.155). Джонкхир (1954) показал, что если хотя бы два из k объемов выборок стремятся к бесконечности так, что n/n_p, n/n_q остаются ограниченными, то распределение U по-прежнему сходится к

нормальному — это можно понять из того, что U есть сумма $\frac{1}{2}k(k-1)$ (не независимых) статистик Вилкоксона U_{pq}. Если r объемов выборок, $r \geqslant 2$, стремятся к бесконечности, то $\frac{1}{2}r(r-1)$ из U_{pq} будут стремиться к нормальности и преобладать в U.

Джонкхир (1954) табулировал точное распределение U для выборок равных объемов m. Его таблица покрывает значения $k = 3$, $m = 2(1)5$; $k = 4$, $m = 2(1)4$; $k = 5$, $m = 2, 3$; $k = 6$; $m = 2$. За этими пределами нормальное приближение удовлетворительно при равных объемах выборок. Даже при очень неравных объемах выборок нормальное приближение кажется удовлетворительным для практических целей, когда $n \geqslant 12$.

Терпстра (1952), предложивший k-выборочный U-критерий и получивший х. ф. и предельное распределение, приведенные выше, дал необходимые и достаточные условия для состоятельности критерия. Если вероятность того, что наблюдение из p-го распределения превосходит наблюдение из q-го, равна (см. (31.51))

$$\mathsf{P}\{x_{pi} > x_{qj}\} = \frac{1}{2} + \varepsilon_{pq}, \quad p \neq q,$$

а взвешенная сумма ε_{pq} равна

$$S_n = \sum_{p<q} \sum n_p n_q \varepsilon_{pq},$$

то условиями состоятельности (когда $n \to \infty$ и n/n_p ограничены при всех p) критерия, использующего большие значения U в качестве критической области, будут: (1) $S_n > 0$, (2) $(kn)^{-3/2} S_n \to \infty$. Они являются непосредственным обобщением условия (31.58) для состоятельности Q при проверке случайности против тренда вниз.

31.74 Насколько нам известно, эффективности k-выборочных критериев H и U не сравнивались. Трудность здесь в том, что формы их предельных распределений различны; при фиксированном k статистика H имеет нецентральное χ^2-распределение, а U — асимптотически нормальное, когда справедлива гипотеза H_0, а также предположительно при общего вида альтернативах. Кажется правдоподобным, что U-критерий будет действовать наилучшим образом, когда альтернативы имеют вид (31.147) с $\theta_1 < \theta_2 < \ldots < \theta_k$, или в более общей ситуации, когда (31.147) заменяется на

$$F_1(x) < F_2(x) < \ldots < F_k(x) \quad \text{при всех} \quad x. \tag{31.156}$$

(31.156) можно назвать *упорядоченной* альтернативной гипотезой. Бартоломью (1961) показал, что критерий U асимптоти-

чески очень эффективен, когда θ_i расположены на равных расстояниях и все n_i равны между собой. Критерий H, с другой стороны, кажется более эффективным против более общих классов альтернатив.

Хогг (1962) дает метод построения свободных от распределения k-выборочных критериев из двухвыборочных критериев.

Критерии симметрии

31.75 Во всех гипотезах, обсуждавшихся в этой главе, мы существенно были связаны с n независимыми наблюдениями (обычно над одной вариантой x, но, в случае проверки независимости, над вектором (x, y)). Наши гипотезы требовали, чтобы некоторые из этих наблюдений были одинаково распределены, и далее проверялись некоторые гипотезы относительно их функций распределения. Мы нашли (см. **31.16**), что для построения подобных критериев для наших гипотез мы должны ограничиться перестановочными критериями, теория распределения которых придает равные вероятности каждому из $n!$ упорядочений наблюдений.

Следствием этой процедуры является то, что полученные нами критерии остаются в силе, если гипотезы, которые мы рассматривали, заменить непосредственно гипотезой, что совместная функция распределения наблюдений инвариантна относительно перестановок ее аргументов. Например, рассмотрим двухвыборочный критерий для гипотезы

$$H_0\colon\ F_1(x) = F_2(x) \quad \text{при всех} \quad x, \tag{31.157}$$

когда имеются случайные выборки объемов n_1 и n_2 соответственно из этих двух распределений и $n = n_1 + n_2$. Обозначим через G совместную функцию распределения n наблюдений. Заменим H_0 гипотезой симметрии

$$H_0'\colon\ G(x_1, x_2, \ldots, x_n) \equiv G(z_1, z_2, \ldots, z_n), \tag{31.158}$$

где $z_1, \ldots, z_n$ — любая перестановка $x_1, \ldots, x_n$. Тогда любой подобный критерий, имеющий заданный размер при гипотезе (31.157), останется таковым и при гипотезе (31.158). Это не значит, что свойства оптимальности критерия будут одними и теми же для обеих гипотез, — обсуждение этого имеется у Лемана и Стейна (1949). Однако это означает, что любой критерий для гипотезы (31.157) не может быть состоятельным против альтернативной гипотезы (31.158).

Практические ситуации, для которых подходят гипотезы симметрии, встречаются часто. Поскольку мы до сих пор не обсуждали этой задачи даже в параметрическом случае, мы начнем с краткого рассмотрения последнего в простейшей ситуации.

Парный t-критерий

31.76 Предположим, что варианты x_1 и x_2 имеют совместно нормальное распределение со средними и дисперсиями (μ_1, σ_1^2), (μ_2, σ_2^2) соответственно и коэффициентом корреляции ρ. Мы желаем проверить сложную гипотезу

$$H_0\colon\ \Delta \equiv \mu_1 - \mu_2 = 0 \qquad (31.159)$$

на основе m независимых наблюдений над (x_1, x_2). Рассмотрим величину $y = x_1 - x_2$. Она распределена нормально, со средним Δ и дисперсией $\sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$. Мы имеем m наблюдений над y и можем, следовательно, проверить H_0 с помощью обычного критерия t Стьюдента для среднего, примененного к разностям $x_{1i} - x_{2i}$, $i = 1, 2, \ldots, m$. Эта процедура применима и при $\rho = 0$, когда x_1 и x_2 — независимые нормальные величины, и в этой ситуации критерий является частным случаем критерия, даваемого формулой (21.51) с $n_1 = n_2 = m$.

31.77 Теперь упростим пример из **31.76**, полагая $\sigma_1^2 = \sigma_2^2$. Совместное распределение $F(x_1, x_2)$ тогда симметрично по x_1 и x_2 с точностью, возможно, до их средних. Когда справедлива гипотеза H_0, мы имеем полную симметрию. Мы можем, следовательно, записать (31.159) как

$$H_0\colon\ F(x_1, x_2) = F(x_2, x_1) \quad \text{при всех} \quad x_1, x_2. \qquad (31.160)$$

Это типичная гипотеза симметрии, которую можно формально представить в общем виде (31.158), записывая G как произведение m сомножителей (по одному для каждого наблюдения над (x_1, x_2)).

Мы теперь отказываемся от нормальной теории пункта **31.76** и ищем свободные от распределения методы проверки непараметрической гипотезы (31.160) для произвольной непрерывной ф. р. F. Если взять, как и раньше, разности $y = x_1 - x_2$, то мы видим, что H_0 влечет симметрию распределения y относительно точки 0, т. е. если G — ф. р. y, то

$$H_0\colon\ G(y) = 1 - G(-y) \quad \text{при} \quad y. \qquad (31.161)$$

Мы, таким образом, свели гипотезу (31.160) о симметрии двумерной ф. р. по ее аргументам к гипотезе (31.161) о симметрии одномерного распределения относительно заданной точки, нуля. Ясно, что эта гипотеза интересна и сама по себе (т. е. нас может просто интересовать симметрия отдельной варианты), и мы переходим к изложению проблемы в этой форме.

31.78 Гипотеза (31.161) означает, что любое наблюденное абсолютное значение y, $|y_i|$, имеет равные вероятности происходить от положительного или отрицательного значения y. Таким образом, имеется 2^n равновероятных выборок, образую-

щихся, если придавать каждому наблюденному $|y_i|$ поочередно знаки плюс и минус и объединять все $|y_i|$ любым возможным способом. Мы создаем основу для построения критерия перестановок, переставляя знаки, приписанные каждому $|y_i|$, и теперь нам нужно выбрать статистику критерия. Если мы рассмотрим альтернативу, что наблюдения нормальны со средним $\theta \neq 0$, то найдем, как в **31.18** и **31.44**, что наиболее мощный критерий перестановок основан на сумме наблюденных y_i (со своими настоящими знаками, конечно). Мы отложим рассмотрение этого критерия (предложенного Фишером (1935a)), поскольку он является частным случаем дисперсионного анализа, который мы будем обсуждать в главе 37, том 3.

31.79 Фишеровский критерий симметрии эквивалентен использованию суммы $|y_i|$ по наблюденным положительным значениям (поскольку сумма всех $|y_i|$ одна и та же для каждой из 2^n перестановок). Для построения рангового критерия симметрии мы можем заменить наблюдения рангами и получить критерий симметрии Вилкоксона, основанный на сумме рангов величин $|y_i|$, взятой по положительным наблюденным значениям, или мы можем использовать эквивалент критерия c_1, основанный на математических ожиданиях порядковых статистик из положительной половины нормального распределения. Эти три критерия симметрии при нормальных альтернативах сдвига имеют по сравнению с t-критерием АОЭ, равные 1, $3/\pi$ и 1 соответственно, точно так же, как и двухвыборочные критерии сдвига. При малых выборках оба ранговых критерия имеют высокую эффективность против нормальных альтернатив (Клоц (1963)); Клоц (1965) рассматривал и другую меру эффективности. Критерий симметрии Вилкоксона не сохраняет своей высокой мощности при некоторых не нормальных альтернативах (с «длинными хвостами»), но он все же лучше, чем критерий Стьюдента, хотя и хуже, чем критерий знаков (Арнольд (1965)), который будет рассмотрен в **32.6**.

Мак Корнак (1965) приводит обширные таблицы критических значений для критерия симметрии Вилкоксона *).

Влияние дискретности: поправки на непрерывность и совпадения

31.80 В ряде случаев как в этой главе, так и в других мы приближали дискретные распределения (в теперешнем контексте — перестановочные распределения статистик критериев) их непрерывными предельными формами. Это приближение часто,

*) Таблица точного распределения статистики критерия симметрии Вилкоксона для $n = 3(1)20$ имеется в *Сборнике статистических таблиц*. (*Прим. ред.*)

хотя и не всегда (см. Плэкетт (1964) и **33.27** ниже) может быть улучшено применением *поправки на непрерывность*, которая сводится к следующему простому правилу: когда последовательные дискретные вероятности в точном распределении связаны с точками z_1, z_2, z_3, то вероятность в точке z_2 считается относящейся к интервалу $\left(\frac{1}{2}(z_1+z_2), \frac{1}{2}(z_2+z_3)\right)$. Таким образом, когда мы хотим найти значение ф. р. в точке z_2 с помощью непрерывной аппроксимации, мы в действительности ищем ее значение в точке $(z_2+z_3)/2$.

31.81 Имеется еще один вопрос, связанный с непрерывностью, который мы должны обсудить здесь. Наши гипотезы относились к наблюдениям из непрерывных ф. р., и это означает, что для любой пары наблюдений вероятность их точного *совпадения* равна нулю и что мы, следовательно, можем пренебречь этой возможностью. Поэтому мы на протяжении этой главы предполагали, что наблюдения могут быть упорядочены без совпадений, так что ранговые статистики были определены однозначно. Однако на практике наблюдения всегда округляются до нескольких значащих цифр, и поэтому совпадения будут иногда появляться. Аналогично, если истинные функции распределения наблюдений на самом деле не непрерывны, а приближенно представлены как непрерывные ф. р., будут появляться совпадения. Как преодолеть эту трудность получения ранжировки при наличии совпадений?

В литературе обсуждались два метода обращения с совпадениями. Первый состоит в том, чтобы упорядочить совпавшие наблюдения случайным образом. Его достоинством является простота, он не требует новой теории, но при этом мы, очевидно, жертвуем информацией, содержащейся в наблюдениях, и можно ожидать, что он будет менее эффективен, чем второй метод, состоящий в том, что каждому из группы совпавших наблюдений приписывается средний ранг этой группы. Достоинства обоих методов исследовались довольно мало, но Путтер (1955) показал, что критерий Вилкоксона при случайном разделении совпадений имеет меньшую АОЭ, чем когда придаются средние ранги. Крускал и Уоллис (1952, 1953) и Крускал (1952) обсуждают совпадения в H-критерии.

До появления дальнейшей информации метод среднего ранга кажется более общеупотребительным. К сожалению, при этом методе пропадает отмечавшееся нами свойство ранговых критериев, что их точные распределения могут быть табулированы раз и навсегда. Потому что если используется метод среднего ранга для разделения совпадений, то на сумму множества рангов это не влияет, но, скажем, их дисперсия изменяется. Перестановочное распределение при малых объемах выборок теперь

зависит от количества и размеров наблюденных совпадений, и это затрудняет табулирование. Кендалл (1962) дает подробные описания и указания относительно необходимых поправок для ранговых коэффициентов корреляции и близких к ним статистик (включая статистику Вилкоксона); другие обсуждения поправок были упомянуты выше.

31.82 Наконец, мы видели в **31.56** и **31.62** и увидим снова в **32.9** и **32.13**, что если распределение наблюдений дискретно, то методы, основанные на предположении непрерывности, оказываются консервативными.

УПРАЖНЕНИЯ

31.1 Пользуясь таблицами распределения χ^2, проверить значения истинных вероятностей в таблице в **31.7.**

31.2 Проверить, что распределение (31.18) имеет моменты (31.12), (31.13) и (31.17).

31.3 Пусть r — коэффициент корреляции, определенный в (31.11), и мы применяем к наблюденным значениям x и y преобразование $X = t_1(x)$, $Y = t_2(y)$ и вычисляем коэффициент корреляции R между преобразованными значениями (X, Y); тогда каждая из $n!$ равновероятных перестановок дает некоторые значения r, R. Показать, что коэффициент корреляции между r и R по $n!$ перестановкам дается формулой

$$C(r, R) = C_1(x, X)\, C_2(y, Y),$$

т. е. что коэффициент корреляции совместного перестановочного распределения наблюденного коэффициента корреляции и коэффициента корреляции преобразованных наблюдений равен просто произведению коэффициента корреляции между x и его преобразованием на коэффициент корреляции между y и его преобразованием.

(Даниэлс, 1944.)

31.4 Вывести четвертый момент рангового коэффициента корреляции, приведенный в (31.22), из общего выражения в (31.14).

31.5 Показать, что (31.21) и (31.40) являются альтернативными определениями (31.19), путем доказательства тождеств (31.20) и (31.39).

31.6 Пользуясь определениями (31.37), (31.38), показать, что коэффициент корреляции в совместном распределении t и r_s по $n!$ равновероятным перестановкам в случае независимости равен $2(n+1)/\{2n(2n+5)\}^{1/2}$.

(См. Даниэлс, 1944.)

31.7 Из непрерывного двумерного распределения получена выборка из n пар (x, y). Пусть $\tilde{x}$ и $\tilde{y}$ — выборочные медианы x и y; определим статистику

$$u = \sum_{i=1}^{n} \operatorname{sign}(x_i - \tilde{x}) \operatorname{sign}(y_i - \tilde{y}).$$

Как можно использовать u для проверки гипотезы независимости x и y? Показать, что его АОЭ по сравнению с выборочным коэффициентом корреляции против альтернатив двумерной нормальности равна $4/\pi^2$. (Это так называемый *критерий срединной корреляции* (medial correlation test).)

(Бломквист, 1950.)

31.8 В **31.36**, пользуясь результатом упражнения 2.15 и симметрией распределения Q, показать, что достаточным условием для того, чтобы критерий случайности, имеющий размер α, был строго несмещенным против альтернатив (31.51), является

$$S_n \geqslant \frac{1}{4} n(n-1) - (1-\alpha)\left\{\frac{1}{2} n(n-1) - Q_0\right\},$$

где Q_0 — критическое значение Q, определенное в (31.55).

(См. Манн, 1945.)

31.9 Показать, что соотношение (31.69) выполняется как для Q, так и для статистики V, определенной в (31.37), и что, следовательно, r_s как критерий случайности имеет ту же АОЭ, что и другой коэффициент ранговой корреляции t.

31.10 В задаче проверки гипотезы случайности (31.49) против альтернатив нормальной регрессии (31.61) рассмотреть класс статистик критериев $S = \sum w_{ij} h_{ij}$, где сумма содержит $n/2$ членов (n кратно 6), причем каждый из индексов принимает $n/2$ различных значений и все индексы различны, а w — веса. Таким образом, в S входят $n/2$ сравнений между независимыми парами наблюдений.

Показать, что S-статистикой с максимальной АОЭ по сравнению с b в (31.63), (31.64) является

$$S_1 = \sum_{k=1}^{n/2} (n - 2k + 1)\, h_{k,\, n-k+1}$$

и ее АОЭ равна

$$A_{S_1,\, b} = (2/\pi)^{1/3} \approx 0{,}86.$$

(Д. Кокс и Стьюарт, 1955.)

31.11 В упражнении 31.10 показать, что если вместо S_1 мы используем форму с равными весами

$$S_2 = \sum_{k=1}^{n/2} h_{k,\, n-k+1},$$

то АОЭ уменьшится до $(3/2\pi)^{1/3} \approx 0{,}78$, но что максимальная АОЭ, достижимая для S-статистики, все веса в которой 1 или 0, равна $(16/9\pi)^{1/3} \approx 0{,}83$, и это есть АОЭ статистики $S_3 = \sum_{k=1}^{n/3} h_{k,\, \frac{2}{3}n+k}$. В S_3 входят только $n/3$ сравнений между «начальными» и «последними» наблюдениями.

(Д. Кокс и Стьюарт, 1955.)

31.12 Определим статистику для проверки случайности $B = \sum_{i=1}^{n/2} \operatorname{sign}(x_i - \tilde{x})$, где $\tilde{x}$ — медиана выборки объема n (n четно). Показать, что ее АОЭ против нормальных альтернатив в точности такова же, как у S_2 в упражнении 31.11.

(Д. Кокс и Стьюарт, 1955; см. Браун и Муд, 1951.)

31.13 Из непрерывного распределения со средним μ и дисперсией σ^2 извлечены независимо N выборок, каждая объема n, и наблюдениям приданы ранги от 1 до n в каждой выборке. Для Nn объединенных вместе наблюде-

ний вычисляется коэффициент корреляции между значениями варианты и соответствующими рангами. Показать, что при $N \to \infty$ он стремится к

$$C_n = \left\{ \frac{12(n-1)}{\sigma^2 (n+1)} \right\}^{1/2} \int_{-\infty}^{\infty} x \left\{ F(x) - \frac{1}{2} \right\} dF(x) = \left(\frac{n-1}{n+1} \right)^{1/2} \frac{3^{1/2} \Delta}{2\sigma},$$

где Δ — коэффициент средней разности Джини, определяемый в (2.24) и упражнении 2.9. В частности, показать, что для нормального распределения

$$C_n = \left\{ \left(\frac{n-1}{n+1} \right) \frac{3}{\pi} \right\}^{1/2},$$

так что

$$C = \lim_{n \to \infty} C_n = (3/\pi)^{1/2} \approx 0{,}98.$$

(Стьюарт, 1954c.)

31.14 Пользуясь (а) теоремой, что корреляция между эффективной оценкой и другой оценкой равна корню квадратному из эффективности последней (см. (17.61)), (б) соотношением между эффективностью оценки и АОЭ, приведенным в **25.13**, (в) теоремой Даниэлса из упражнения 31.3 и (г) последним результатом из упражнения 31.13, получить результаты об АОЭ рангового коэффициента корреляции r_s (а следовательно, и t) как критерия независимости (31.48) и как критерия случайности (31.70), а также получить АОЭ критерия суммы рангов Вилкоксона против нормальных альтернатив (31.117).

(Стьюарт, 1954c.)

31.15 Получить дисперсию статистики Вилкоксона, даваемую в (31.105), рассматривая среднее выборки из n_1 целых чисел, извлеченной из конечной генеральной совокупности, образованной первыми n натуральными числами.

(Крускал и Уоллис, 1952, 1953.)

31.16 Для двухвыборочной статистики Вилкоксона U в **31.56** показать, что, каковы бы ни были распределения наблюдений F_1, F_2,

$$\mathbf{M}(U) = n_1 n_2 p, \qquad \mathbf{D}U = O(N^3),$$

где N означает любое из n_1, n_2, n. Вывести отсюда, что, когда $n_1, n_2 \to \infty$ с фиксированным n_1/n_2, критерий состоятелен, $p \neq 1/2$.

(Питмэн, 1948.)

31.17 Показать, что для логистического распределения из упражнения 17.5 АОЭ критерия Вилкоксона по сравнению с критерием t Стьюдента для сдвига равна $\pi^2/9$.

(Это вместе с результатами упражнения 17.5 и **25.13** означает, что критерий Вилкоксона является асимптотически наилучшим. Кейпон (1961) показал, что это справедливо в общем случае для локально наиболее мощных ранговых критериев, каковым критерий Вилкоксона является в данном случае.)

31.18 Для распределения

$$dF = \exp(-x)\, x^{p-1}\, dx/\Gamma(p), \qquad 0 \leqslant x < \infty, \quad p > 1/2,$$

показать, что АОЭ критерия Вилкоксона по сравнению с критерием t Стьюдента для различия в сдвиге между двумя выборками равна

$$A_{U,t} = \frac{3p}{2^{4(p-1)} \{(2p-1)\, \mathrm{B}(p, p)\}^2}$$

и это монотонно убывающая функция от p. Проверить, что $A_{U,t} > 1{,}25$ при $p \leqslant 3$. Показать, что при $p \to 1/2$ $A_{U,t} \to \infty$ и что при $p \to \infty$ $A_{U,t} \to 3/\pi$ в согласии с (31.117).

31.19 Показать, что H-критерий из **31.71** при $k = 2$ сводится к критерию Вилкоксона с критической областью, поровну распределенной между обоими «хвостами» статистики критерия.

31.20 Пользуясь результатом из **25.15** об АОЭ двух статистик критериев с предельными нецентральными χ^2-распределениями, имеющими одинаковое число степеней свободы и параметры нецентральности, зависящие от расстояния от H_0, установить, что k-выборочный H-критерий из **31.71** имеет по сравнению со стандартным F-критерием в нормальном случае АОЭ, равную (31.115).

(См. Эндрюс, 1954.)

31.21 Показать, что при проверке гипотезы $\rho = 0$ для двумерной нормальной генеральной совокупности выборочный коэффициент корреляции r дает РНМН критерии против одно- и двусторонних альтернатив.

31.22 Показать, что критерий Вилкоксона имеет по сравнению с критерием t Стьюдента АОЭ, равную 1, против альтернатив сдвига для прямоугольных распределений.

(Питмэн, 1948.)

31.23 Пользуясь (31.115), (31.136) и (31.138), показать, что АОЭ критерия Вилкоксона по сравнению с критерием c_1 для альтернатив сдвига равна

$$A_{U,c_1} = 12\left[\int \{f_1(x)\}^2\,dx \Big/ \int \frac{\{f_1(x)\}^2\,dx}{\varphi\{J[F_1(x)]\}}\right]^2$$

и что, следовательно,

$$0 \leqslant A_{U,c_1} \leqslant 6/\pi.$$

(Ходжес и Леман (1961) показали, что оба равенства достижимы.)

31.24 Показать, что интервалы (31.95), основанные на двухвыборочной статистике Вилкоксона U, определяемой в (31.98), равны $\{D(U_\alpha), D(n_1n_2 + 1 - U_\alpha)\}$, где $\mathsf{P}\{U < U_\alpha \mid H_0\} = \alpha/2$, а $D(r)$ есть r-я наименьшая из n_1n_2 разностей $(x_{1i} - x_{2j})$.

(Леман, 1963.)

31.25 Показать, что критерий Вилкоксона для симметрии, рассмотренный в **31.79**, дает для параметра симметричного распределения свободные от распределения доверительные интервалы $\left\{A(W_\alpha), A\left(\frac{1}{2}n(n-1) + 1 - W_\alpha\right)\right\}$, где W_α — нижнее критическое значение критерия с «равными хвостами», а $A(r)$ есть r-е наименьшее из $n(n-1)/2$ средних $(x_i + x_j)/2$, $i < j = 1, 2, \ldots, n$.

(Леман, 1963.)

ГЛАВА 32

НЕКОТОРЫЕ ПРИМЕНЕНИЯ ПОРЯДКОВЫХ СТАТИСТИК

32.1 В главе 31 мы нашли, что простые и чрезвычайно эффективные критерии перестановок для некоторых непараметрических гипотез могут быть получены путем использования рангов, отражающих соотношения порядка между наблюдениями. В настоящей главе мы сначала обсуждаем использование самих порядковых статистик для построения свободных от распределения процедур для непараметрических задач (5) и (6), перечисленных в **31.13**. Затем мы рассматриваем применения порядковых статистик в других (параметрических) ситуациях. Напомним читателю, что общая теория распределений порядковых статистик была рассмотрена в главе 14 и что теория несмещенного оценивания с минимальной дисперсией параметров сдвига и масштаба посредством линейных функций от порядковых статистик была дана в главе 19. Очень полезный общий обзор литературы по порядковым статистикам был дан Уилксом (1948), обширную библиографию которого дополняет более поздняя библиография Ф. Дэвида и Джонсона (1956). Изложение многих разделов теории с обширными таблицами дается у Сархана и Гринберга (1962). Обзор теории «выборочных промежутков» (разностей между последовательными порядковыми статистиками) дается Пайком (1965).

Критерий знаков для квантилей

32.2 Так называемый критерий знаков для значения квантили непрерывного распределения был, по-видимому, первым из когда-либо использовавшихся свободных от распределения критериев*), но современный интерес к нему берет начало с работы Кокрэна (1937).

---

*) Тодхантер (1865) ссылается на его использование в простой форме Джоном Арбетнотом (врач королевы Анны, а ранее учитель математики) для обоснования «Аргумента в пользу божественного провидения, взятого из постоянной регулярности, наблюдаемой в рождении обоих полов» (1710—1712); Арбетнот был известен своим остроумием, им написана сатира «Искусство политической лжи».

Предположим, что ф. р. наблюдений есть $F(x)$ и что

$$F(X_p) = p, \tag{32.1}$$

так что X_p есть p-квантиль этого распределения, т. е. значение, ниже которого лежит $100p$ процентов распределения. Для любого p, $0 < p < 1$, значение X_p есть характеристика положения распределения. Мы хотим проверить гипотезу

$$H_0\colon\ X_p = x_0, \tag{32.2}$$

где x_0 — некоторое заданное значение. (Если для удобства принять x_0 за начало отсчета, то мы хотим проверить равенство X_p нулю.)

32.3 Если имеется выборка из n наблюдений, то мы знаем что выборочная функция распределения будет сходиться по вероятности к ф. р. наблюдений. Отметим соотношение между порядковыми статистиками $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ и гипотетическим значением X_p, подлежащим проверке. Сосчитаем, сколько наблюдений в выборке попадает ниже x_0, т. е. образуем статистику

$$S = \sum_{i=1}^{n} h(x_0 - x_{(i)}) = \sum_{i=1}^{n} h(x_0 - x_i), \tag{32.3}$$

где (см. (31.36))

$$h(z) = \begin{cases} 1, & z > 0, \\ 0, & z < 0. \end{cases}$$

Статистика S считает число положительных значений среди разностей $(x_0 - x_i)$, и поэтому критерий, основанный на S, называется критерием знаков*). Сразу видно, что S имеет биномиальное распределение, поскольку S есть сумма n независимых наблюдений над $(0-1)$-случайной величиной $h(x_0 - x)$ с

$$\mathsf{P}\{h(x_0 - x) = 1\} = \mathsf{P}\{x < x_0\}.$$

Обозначим $\mathsf{P}\{x < x_0\} = P$. Гипотеза (32.2) сводится к

$$H_0\colon\ P = p, \tag{32.4}$$

и мы просто проверяем гипотезу о биномиальном параметре P. Нас могут интересовать односторонние или двусторонние альтернативы к (32.4).

Если мы больше ничего не знаем относительно ф. р. $F(x)$, то интуитивно очевидно, что мы не можем получить ничего луч-

*) Вследствие непрерывности ф. р. наблюдений событие $x_i - x_0 = 0$ может произойти лишь с вероятностью нуль. Если такие «совпадения» происходят на практике, наиболее мощная процедура состоит в том, чтобы игнорировать эти наблюдения при проверке гипотезы, как было показано Хемелрейком (1952) (см. **31.81**).

шего, чем S, в качестве статистики критерия, и мы находим из биномиальной теории (см. упражнение 22.2 и **23.31**), что для односторонней альтернативы $H_1: P > p$ критическая область, состоящая из больших значений S, является РНМ, а для двусторонней альтернативы $H_2: P \neq p$ двусторонняя критическая область является РНМН.

В наиболее важном для практики случае, когда $p = 1/2$ и мы проверяем медиану распределения, мы имеем для S симметричное биномиальное распределение и РНМН критическая область против H_2 симметрична.

Формальное доказательство этих результатов дано в книге Лемана (1959).

32.4 Таким образом, при малом объеме выборки n таблицы биномиального распределения достаточны как для определения размера критерия знаков, так и для определения его мощности против любого конкретного альтернативного значения P, а следовательно, и его функции мощности для альтернатив H_1 или H_2. Когда n возрастает, сходимость биномиального распределения к нормальному позволяет нам сказать, что $(S - nP)/\{nP(1-P)\}^{1/2}$ имеет стандартное нормальное распределение. Если мы пользуемся поправкой на непрерывность, как в **31.80**, то это сводится к замене $|S - nP|$ на $|S - nP| - \frac{1}{2}$ при выполнении критерия.

В случае медианы, когда мы проверяем гипотезу $P = 1/2$, сходимость к нормальности настолько быстрая, что здесь, скорее всего, вовсе не потребуются специальные таблицы, поскольку нам нужно только сравнить значение

$$\left(\left|S - \tfrac{1}{2}n\right| - \tfrac{1}{2}\right) \Big/ \left(\tfrac{1}{2}n^{1/2}\right) \qquad (32.5)$$

с подходящим стандартным нормальным отклонением. Кокрэн (1937) дает точные критические значения для $n \leqslant 50$ и размера критерия $\alpha = 0{,}05$, Диксон и Муд (1946) — для $n \leqslant 100$ и $\alpha = 0{,}25$, 0,1, 0,05 и 0,01, и Маккиннон (1964) — для $n = 1(1)1000$ и $\alpha = 0{,}50$, 0,10, 0,05, 0,02, 0,01 и 0,001 *.

Мощность критерия знаков для медианы

32.5 Приближенная мощность критерия знаков также легко устанавливается с помощью нормальной аппроксимации. Пренебрегая поправкой на непрерывность, указанной в **32.4**, поскольку она мала при больших выборках, мы видим, что критической областью одностороннего критерия для $P = 1/2$ против

*) Таблицы критических значений имеются также в *Таблицах математической статистики* и *Сборнике статистических таблиц*. (*Прим. ред.*)

$P > 1/2$ является

$$S \geqslant \tfrac{1}{2} n + d_\alpha \tfrac{1}{2} n^{1/2},$$

где d_α — подходящее нормальное отклонение для критерия размера α. Следовательно, функция мощности приближенно равна

$$Q_1(P) = \int\limits_{\frac{1}{2}n+\frac{1}{2}d_\alpha n^{1/2}}^{\infty} \{2\pi nP(1-P)\}^{-1/2} \exp\left\{-\frac{1}{2}\frac{(u-nP)^2}{nP(1-P)}\right\} du =$$

$$= \int\limits_{\frac{n^{1/2}\left(\frac{1}{2}-P\right)+\frac{1}{2}d_\alpha}{\{P(1-P)\}^{1/2}}}^{\infty} (2\pi)^{-1/2} \exp\left(-\frac{1}{2}t^2\right) dt = G\left\{\frac{n^{1/2}\left(P-\frac{1}{2}\right)-\frac{1}{2}d_\alpha}{[P(1-P)]^{1/2}}\right\}, \tag{32.6}$$

где $G\{x\}$ — нормальная ф. р. Из (32.6) сразу следует, что при $n \to \infty$ мощность $\to 1$ для любого $P > 1/2$, так что критерий состоятелен. Аналогично находим, что функция мощности двустороннего симметричного критерия с критической областью

$$\left|S - \tfrac{1}{2}n\right| \geqslant d_{\alpha/2} \cdot \tfrac{1}{2} n^{1/2}$$

равна

$$Q_2(P) = G\left\{\frac{n^{1/2}\left(P-\frac{1}{2}\right)-\frac{1}{2}d_{\alpha/2}}{[P(1-P)]^{1/2}}\right\} + G\left\{\frac{n^{1/2}\left(\frac{1}{2}-P\right)-\frac{1}{2}d_{\alpha/2}}{[P(1-P)]^{1/2}}\right\} \tag{32.7}$$

и стремится к 1 для любого $P \neq 1/2$ при $n \to \infty$. Этим устанавливается состоятельность двустороннего критерия против общих альтернатив.

Диксон (1953b) табулировал мощность двустороннего критерия знаков для размеров критерия $\alpha \leqslant 0{,}05$, $\alpha \leqslant 0{,}01$, n, изменяющегося от 5 до 100, и $P = 0{,}05(0{,}05)0{,}95$. Мак Стьюарт (1941) дает таблицу минимального объема выборки, который требуется для достижения заданной мощности против данного значения *). Гиббонс (1964) исследовал влияние отклонения от нормальности на мощность одностороннего критерия знаков.

Критерий знаков в симметричном случае

32.6 Функции мощности (32.6) и (32.7) выражаются через значение P, задаваемое альтернативной гипотезой. Если мы теперь желаем рассматривать эффективность критерия знаков в конкретных ситуациях, нам нужна дальнейшая конкретизация

*) Эта таблица приведена в *Сборнике статистических таблиц*. (*Прим. ред.*)

распределения. Если мы вернемся к первоначальной формулировке гипотезы (32.2) и ограничимся случаем медианы $X_{0,5}$, которую будем обозначать M, то наша задача состоит в проверке гипотезы

$$H_0\colon\ M = M_0. \tag{32.8}$$

Если функция распределения наблюдений, как и раньше, равна $F(x)$, а ф. п. равна $f(x)$, то мы имеем для значения P

$$P = F(M_0) = \int_{-\infty}^{M_0} f(x)\,dx. \tag{32.9}$$

Предположим, что нас интересует относительная эффективность критерия знаков, когда известно, что ф. р. F симметрична, так что ее среднее и медиана M совпадают. Мы можем в этой ситуации проверять гипотезу (32.8), пользуясь в качестве статистики критерия выборочным средним $\bar{x}$. Если F имеет конечную дисперсию σ^2, то $\bar{x}$ асимптотически нормально со средним M и дисперсией σ^2/n, и при больших выборках оно эквивалентно статистике Стьюдента

$$t = \frac{\bar{x} - M}{s/(n-1)^{1/2}},$$

где s^2 — выборочная дисперсия. Для $\bar{x}$ мы имеем

$$\frac{\partial}{\partial M}\,\mathsf{M}(\bar{x} \mid M) = 1, \quad \mathsf{D}(\bar{x} \mid M) = \sigma^2/n,$$

так что

$$\frac{\left\{\frac{\partial}{\partial M}\,\mathsf{M}(\bar{x} \mid M)\right\}^2}{\mathsf{D}(\bar{x} \mid M)} = \frac{n}{\sigma^2}. \tag{32.10}$$

Для статистики критерия знаков нам будет удобно принять M за начало отсчета, так что

$$\mathsf{M}(S \mid P) = nP = n\int_{-\infty}^{M_0 - M} f(x)\,dx$$

и, следовательно,

$$\left\{\frac{\partial \mathsf{M}(S \mid M)}{\partial M}\right\}_{M=M_0} = -nf(0).$$

Кроме того,

$$\mathsf{D}(S \mid M_0) = \frac{1}{4}n.$$

Таким образом, в первоначальной системе отсчета

$$\frac{\left\{\frac{\partial \mathsf{M}(S \mid M)}{\partial M}\right\}^2_{M=M_0}}{\mathsf{D}(S \mid M_0)} = 4n\{f(M)\}^2. \tag{32.11}$$

Из (32.10), (32.11) и (25.27) находим, что эффективность критерия знаков равна

$$A_{S,\bar{x}} = 4\sigma^2\{f(M)\}^2 \tag{32.12}$$

— результат, принадлежащий Питмэну.

32.7 Ясно, что (32.12) не имеет ненулевой нижней границы, как это было для АОЭ критериев Вилкоксона и Фишера — Иэйтса в главе 31, поскольку мы можем иметь для медианной ординаты $f(M) = 0$. В нормальном случае $f(M) = (2\pi\sigma^2)^{-1/2}$, так что (32.12) принимает значение $2/\pi$. Поскольку мы проверяем здесь симметрию относительно M_0, мы можем, как указано в **31.78**, использовать критерий Вилкоксона, имеющий АОЭ, равную $3/\pi$ в нормальном случае и всегда превышающую 0,864. Таким образом, имеется мало оснований, за исключением простоты, чтобы рекомендовать пользоваться критерием знаков в качестве критерия симметрии относительно заданной медианы: в такой ситуации эффективнее проверять выборочное среднее. Критерий знаков полезен, когда мы хотим проверять гипотезу о медиане без предположения симметрии.

Диксон (1953b) табулирует эффективность мощности двустороннего критерия знаков в нормальном случае (и дает ссылки на ранние работы, в частности, Уолша). Он показывает, что относительная эффективность (т. е. обратная величина отношения объемов выборок, требуемых для критерия знаков и критерия t Стьюдента для достижения одинаковой мощности при равных размерах критериев и против одних и тех же альтернатив, — см. **25.2**) убывает при возрастании объема выборки, размера критерия или расстояния P от 1/2.

Уиттинг (1960) использует разложение Эджворта до порядка n^{-2} и показывает, что эта аппроксимация второго порядка дает результаты для АОЭ, мало отличающиеся от (32.12).

Свободные от распределения доверительные интервалы для квантилей

32.8 Совместное распределение порядковых статистик непосредственно зависит от ф. р. наблюдений (см., например, (14.1) и (14.2)), и, следовательно, точечная оценка квантилей распределения с помощью порядковых статистик не является свободной от распределения. Замечательно, однако, что пары порядковых статистик могут быть использованы для построения свободных от распределения доверительных интервалов для любой квантили распределения.

Рассмотрим пару порядковых статистик $x_{(r)}$ и $x_{(s)}$, $r < s$, в выборке из n наблюдений из непрерывной ф. р. $F(x)$. Формула (14.2) дает для совместного распределения $F_r = F(x_{(r)})$ и $F_s = F(x_{(s)})$ выражение

$$dG_{r,s} = \frac{F_r^{r-1}(F_s - F_r)^{s-r-1}(1 - F_s)^{n-s}\,dF_r\,dF_s}{\mathrm{B}(r, s-r)\,\mathrm{B}(s, n-s+1)}. \tag{32.13}$$

p-квантиль X_p распределения $F(x)$ определяется равенством (32.1). Интервал $(x_{(r)}, x_{(s)})$ может накрыть X_p, только если $F_r \leqslant p \leqslant F_s$, и вероятность этого события равна

$$1 - \alpha = \int_0^p \int_p^1 dG_{r,s},$$

где первый интеграл относится к F_r. Это равно

$$1 - \alpha = \int_0^p \int_0^1 dG_{r,s} - \int_0^p \int_0^p dG_{r,s}, \tag{32.14}$$

и, поскольку $F_r \leqslant F_s$, (32.14) можно переписать как

$$1 - \alpha = \int_0^p \int_0^1 dG_{r,s} - \int_0^{F_s} \int_0^p dG_{r,s}. \tag{32.15}$$

Значения двойных интегралов в (32.15) легко найти. В первом из них интегрирование по F_s охватывает целиком его область изменения, и, следовательно, интеграл от 1 до p берется от маргинального распределения F_r, которое согласно (14.1) равно

$$dG_r = \frac{F_r^{r-1}(1 - F_r)^{n-r}\, dF_r}{\mathrm{B}(r,\ n - r + 1)};$$

это бета-распределение первого рода, ф. р. которого есть просто неполная бета-функция. Следовательно,

$$\int_0^p \int_0^1 dG_{r,s} = \int_0^p dG_r = I_p(r,\ n - r + 1). \tag{32.16}$$

Во втором двойном интеграле в (32.15) мы делаем подстановку $u = F_r/F_s$, $v = F_s$ с якобианом v точно так же, как в **11.9**, и находим

$$\int_0^{F_s} \int_0^p dG_{r,s} = \int_0^p \left\{ \int_0^1 \frac{(uv)^{r-1}(v - uv)^{s-r-1}(1 - v)^{n-s}}{\mathrm{B}(r,\ s - r)\,\mathrm{B}(s,\ n - s + 1)}\, du \right\} v\, dv,$$

и после интегрирования по u по всей его области от 0 до 1 нам, как и прежде, остается проинтегрировать маргинальное распределение, на этот раз F_s, от 0 до p. Таким образом,

$$\int_0^{F_s} \int_0^p dG_{r,s} = \int_0^p dG_s = I_p(s,\ n - s + 1). \tag{32.17}$$

Подставляя (32.16), (32.17) в (32.15), получаем

$$\mathbf{P}\{x_{(r)} \leqslant X_p \leqslant x_{(s)}\} = 1 - \alpha = I_p(r, n-r+1) - I_p(s, n-s+1). \tag{32.18}$$

32.9 Мы видим из (32.18), что интервал $(x_{(r)}, x_{(s)})$ покрывает квантиль X_p с коэффициентом доверия, который вовсе не зависит от $F(x)$, и мы, таким образом, имеем свободный от распределения доверительный интервал для X_p. Поскольку $I_p(a, b) = 1 - I_{1-p}(b, a)$, мы можем также записать коэффициент доверия как

$$1 - \alpha = I_{1-p}(n-s+1, s) - I_{1-p}(n-r+1, r). \tag{32.19}$$

Вследствие связи неполной бета-функции с биномиальным разложением, приведенной в **5.7**, (32.18) можно выразить как

$$1 - \alpha = \left\{ \sum_{i=0}^{s-1} - \sum_{i=0}^{r-1} \right\} \binom{n}{i} p^i q^{n-i} = \sum_{i=r}^{s-1} \binom{n}{i} p^i q^{n-i}, \tag{32.20}$$

где $q = 1 - p$. Коэффициент доверия равен, следовательно, сумме членов бинома $(q+p)^n$ от $(r+1)$-го до s-го включительно.

Если выбирать пару симметрично расположенных порядковых статистик, то $s = n - r + 1$ и мы получаем в (32.18)—(32.20)

$$1 - \alpha = I_p(r, n-r+1) - I_p(n-r+1, r) =$$
$$= 1 - \{I_{1-p}(n-r+1, r) + I_p(n-r+1, r)\} = \tag{32.21}$$
$$= \sum_{i=r}^{n-r} \binom{n}{i} p^i q^{n-i}, \tag{32.22}$$

так что коэффициент доверия есть сумма $n - 2r + 1$ центральных членов биномиального разложения, в котором по r членов с каждого конца опущены.

Для любых значений r и n коэффициент доверия, отвечающий интервалу $(x_{(r)}, x_{(n-r+1)})$, можно найти из (32.21), (32.22), пользуясь, если необходимо, таблицами *Tables of the Incomplete Beta Function*. Могут также быть использованы таблицы биномиального распределения, указанные в **5.7**. Упражнение 32.4 предоставляет читателю возможность попрактиковаться в вычислениях.

Шеффе и Тьюки (1945) показали, что если распределение наблюдений дискретно, то указанные выше доверительные интервалы покрывают X_p с вероятностью $\geqslant 1 - \alpha$.

32.10 В частном случае медианы распределения $X_{0,5}$ формулы (32.21), (32.22) принимают особенно простой вид:

$$1-\alpha=1-2I_{0,5}(n-r+1,\ r)=2^{-n}\sum_{i=r}^{n-r}\binom{n}{i}. \qquad (32.23)$$

Эта процедура построения доверительного интервала для медианы была впервые предложена Томпсоном (1936).

Маккиннон (1964) дает таблицы для $n=1(1)1000$ и α, насколько возможно близких к 0,50, 0,10, 0,05, 0,02, 0,01 и 0,001. Нэйр (1940) дает аналогичные таблицы для $n=6(1)81$, указывая точное значение α, которое бралось насколько возможно близким к 0,05 и 0,01 *).

Свободные от распределения толерантные интервалы

32.11 В **20.37** мы рассматривали задачу нахождения толерантных границ для нормальной ф. р. Предположим теперь, что нам нужно строить такие интервалы, не делая никаких предположений о форме распределения, помимо непрерывности. Нам требуется найти случайный интервал (l, u) такой, что

$$\mathbf{P}\left\{\int_l^u f(x)\,dx\geqslant\gamma\right\}=\beta, \qquad (32.24)$$

где $f(x)$ — неизвестная непрерывная функция плотности. Не очевидно, что такая процедура, свободная от распределения, возможна, но Уилкс (1941, 1942) показал, что порядковые статистики $x_{(r)}$, $x_{(s)}$ дают свободные от распределения толерантные границы, а Роббинс (1944) показал, что это могут быть только порядковые статистики.

Если положить $l=x_{(r)}$, $u=x_{(s)}$ в соотношении (32.24), то мы можем переписать его в виде

$$\mathbf{P}[\{F(x_{(s)})-F(x_{(r)})\}\geqslant\gamma]=\beta. \qquad (32.25)$$

Мы можем получить точное распределение случайной величины $F(x_{(s)})-F(x_{(r)})$ из (32.13) с помощью преобразования $y=F(x_{(s)})-F(x_{(r)})$, $z=F(x_{(r)})$, имеющего якобиан 1. Тогда (32.13) принимает вид

$$dH_{y,z}=\frac{z^{r-1}y^{s-r-1}(1-y-z)^{n-s}\,dy\,dz}{\mathrm{B}(r,\ s-r)\,\mathrm{B}(s,\ n-s+1)},\quad 0\leqslant y+z\leqslant 1. \qquad (32.26)$$

*) Для построения доверительного интервала для медианы могут быть использованы таблицы критических значений критерия знаков, см. **32.4**. (*Прим. ред.*)

Проинтегрировав (32.26) по z по всей его области изменения $(0, 1-y)$, получим для маргинального распределения y выражение

$$dG_{s-r} = \frac{y^{s-r-1}\,dy}{\mathrm{B}(r, s-r)\,\mathrm{B}(s, n-s+1)} \int_0^{1-y} z^{r-1}(1-y-z)^{n-s}\,dz. \quad (32.27)$$

Положим $z = (1-y)t$; тогда (32.27) преобразуется в

$$dG_{s-r} = \frac{y^{s-r-1}(1-y)^{n-s+r}\,dy}{\mathrm{B}(r, s-r)\,\mathrm{B}(s, n-s+1)} \int_0^1 t^{r-1}(1-t)^{n-s}\,dt =$$

$$= y^{s-r-1}(1-y)^{n-s+r}\,dy\,\frac{\mathrm{B}(r, n-s+1)}{\mathrm{B}(r, s-r)\,\mathrm{B}(s, n-s+1)} =$$

$$= \frac{y^{s-r-1}(1-y)^{n-s+r}\,dy}{\mathrm{B}(s-r, n-s+r+1)}, \quad 0 \leqslant y \leqslant 1. \quad (32.28)$$

Таким образом, $y = F(x_{(s)}) - F(x_{(r)})$ имеет бета-распределение первого рода. Если мы положим $r = 0$ в (32.28), а $F(x_{(0)})$ примем равным нулю (так что $x_{(0)} = -\infty$), то (32.28) сведется к (14.1), где вместо r написано s.

32.12 Из (32.28) мы видим, что (32.25) приводится к виду

$$\mathsf{P}\{y \geqslant \gamma\} = \int_\gamma^1 \frac{y^{s-r-1}(1-y)^{n-s+r}\,dy}{\mathrm{B}(s-r, n-s+r+1)} = \beta, \quad (32.29)$$

что в терминах неполной бета-функции можно переписать как

$$\mathsf{P}\{F(x_{(s)}) - F(x_{(r)}) \geqslant \gamma\} = 1 - I_\gamma(s-r, n-s+r+1) = \beta. \quad (32.30)$$

Соотношение (32.30) для свободного от распределения толерантного интервала $(x_{(r)}, x_{(s)})$ содержит пять величин: γ (минимальная доля $F(x)$, которую мы хотим накрыть), β (вероятность, с которой мы желаем это сделать), объем выборки n и положения порядковых статистик в выборке, r и s. Если заданы любые четыре из них, мы можем решить (32.20) относительно пятой. На практике β и γ обычно фиксированы как уровни, требуемые постановкой задачи, а r и s выбираются симметрично, так что $s = n - r + 1$. Тогда (32.30) сводится к

$$I_\gamma(n-2r+1, 2r) = 1 - \beta. \quad (32.31)$$

Левая часть (32.31) — монотонно возрастающая функция от n, и для любых фиксированных β, γ, r мы можем выбрать n достаточно большим, чтобы выполнялось (32.31). На практике мы должны выбирать n как ближайшее целое число, превосхо-

дящее решение (32.31). Если $r = 1$, так что используются экстремальные значения выборки, (32.31) сводится к

$$I_\gamma(n - 1{,}2) = 1 - \beta; \tag{32.32}$$

этим определяется вероятность β, с которой размах выборки из n наблюдений покрывает по меньшей мере долю γ распределения, из которого получена выборка.

Решение (32.20) (и его частных случаев (32.31), (32.32)) должно проводиться численно с помощью таблиц *Tables of the Incomplete Beta Function* или, что равносильно, таблиц биномиальной ф. р. (см. **5.7**). Мёрфи (1948) дает графики γ как функции от n для $\beta = 0{,}90$, $0{,}95$ и $0{,}99$ и $r + (n - s + 1) = 1(1)6(2)10(5)30(10)60(20)100$; они являются точными при $n \leqslant 100$ и приближенными до $n = 500$.

Пример 32.1

Рассмотрим численное решение уравнения (32.32) относительно n. Его можно переписать как

$$1 - \beta = \frac{1}{\mathrm{B}(n - 1{,}2)} \int_0^\gamma y^{n-2}(1 - y)\,dy = n(n - 1)\left\{\frac{\gamma^{n-1}}{n - 1} - \frac{\gamma^n}{n}\right\} = n\gamma^{n-1} - (n - 1)\gamma^n. \tag{32.33}$$

Для значений β, γ, требуемых на практике (обычно 0,90 или больше), n настолько велико, что можно записать (32.33) приближенно как

$$1 - \beta = n\gamma^{n-1}(1 - \gamma),$$

$$\gamma = \left\{\left(\frac{1 - \beta}{1 - \gamma}\right)\frac{1}{n}\right\}^{1/(n-1)} \tag{32.34}$$

или

$$\log n + (n - 1)\log\gamma = \log\{(1 - \beta)/(1 - \gamma)\}. \tag{32.35}$$

Производная левой части (32.35) по n равна $(1/n) + \log\gamma$, и при больших n левая часть (32.35) монотонно убывает по n. Поэтому мы можем взять какое-нибудь пробное значение n, сравнить левую часть с (фиксированной) правой частью (32.35) и увеличить (уменьшить) n, если левая (правая) часть больше. Значение n, удовлетворяющее приближенно (32.35), будет больше того, которое удовлетворяет точному соотношению (32.33), поскольку из правой части последнего был выброшен положительный член γ^n, и мы безопасно можем использовать (32.35) без уточнений, либо можем подставить (32.35) в (32.33) и найти поправку, чтобы получить точное значение.

Пример 32.2

Проиллюстрируем пример 32.1 конкретным вычислением. Положим $\beta = \gamma = 0{,}99$. Тогда (32.35) переходит в

$$\log n + (n-1)\log 0{,}99 = 0,$$

поскольку правая часть, очевидно, равна нулю при $\beta = \gamma$. Мы можем пользоваться логарифмами по основанию 10, так как переход к натуральным логарифмам не меняет (32.35). Таким образом, нам нужно решить уравнение

$$\log_{10} n - 0{,}00436\,(n-1) = 0.$$

Начнем с $n = 1000$. При этом значении левая часть отрицательна, поэтому мы уменьшаем n до 500, тогда она становится положительной. Мы далее продвигаемся следующим образом.

n	$\log_{10} n$	$0{,}00436\,(n-1)$	n	$\log_{10} n$	$0{,}00436\,(n-1)$
1000	3	4,36	600	2,7782	2,61
500	2,6990	2,18	640	2,8062	2,79
700	2,8451	3,05	645	2,8096	2,81
650	2,8129	2,83			

Подставим теперь значение $n = 645$ в точное уравнение (32.33). Его правая часть равна

$$645\,(0{,}99)^{644} - 644\,(0{,}99)^{645} = 1{,}004 - 0{,}992 = 0{,}012.$$

Левая часть этого равенства есть $1 - \beta = 0{,}01$, так что согласие хорошее, и мы для всех практических целей можем брать $n = 645$, чтобы получить 99-процентный толерантный интервал для 99 процентов исходной ф. р.

32.13 Мы рассмотрели только простейший случай построения свободных от распределения толерантных интервалов для одномерного непрерывного распределения. Распространение на многомерные толерантные области, включая разрывный случай, было сделано в работах Вальда (1943b), Шеффе и Тьюки (1945), Тьюки (1947, 1948), Фрэзера и Уормлитона (1951), Фрэзера (1951, 1953) и Кемпермана (1956). Уилкс (1948) дает обзор результатов, имевшихся к тому времени. Уолш (1962) рассматривает симметричные непрерывные распределения. Гудмэн и Мадански (1962) рассмотрели свободные от параметра и свободные от распределения толерантные пределы для экспоненциального распределения.

Шеффе и Тьюки (1945) и Тьюки (1948) показали, что если исходное распределение дискретно, то приведенные выше толерантные интервалы и области имеют вероятность $\geqslant 1 - \alpha$.

Точечное оценивание с помощью порядковых статистик

32.14 Как мы отмечали в начале **31.8**, мы не можем получить свободных от распределения точечных оценок, пользуясь порядковыми статистиками, так как их совместное распределение существенно зависит от ф. р. $F(x)$. Поэтому мы снова вступаем в область параметрических задач и ставим вопрос, как можно использовать порядковые статистики в оценке параметров. Имеются две существенно различные ситуации, в которых могут рассматриваться порядковые статистики:

(1) Мы можем сознательно использовать функции от порядковых статистик для оценки параметров, зная даже, что эти процедуры оценки неэффективны, ввиду простоты и скорости вычислений. (По существу, тот же вопрос в другой связи мы обсуждали в примере 17.13.) В **14.6—7** мы приводили некоторые числовые данные, касающиеся эффективностей выборочной медианы и средины размаха как оценки среднего значения нормальной совокупности, а также оценки для стандартного отклонения нормальной совокупности, основанной на выборочной интерквартильной широте. Эти три оценки служат примерами легко вычисляемых неэффективных статистик.

(2) По тем или иным причинам не все элементы выборки могут быть в наличии при оценивании, и мы вынуждены использовать оценку, являющуюся функцией лишь некоторых из них. Таким образом, различие между (1) и (2) относится, по существу, к происхождению задачи. Формально, однако, мы можем считать (1) крайним случаем (2), когда число недостающих элементов выборки равно нулю.

Урезание и цензурирование

32.15 Прежде чем переходить к каким-либо деталям, мы кратко рассмотрим обстоятельства, при которых могут отсутствовать элементы выборки. Предположим сначала, что основная варианта x просто не может быть наблюдаема в какой-нибудь части или частях ее области значений. Например, если x есть расстояние от центра вертикальной круговой мишени фиксированного радиуса R, то мы можем наблюдать x только для тех выстрелов, которые попали в мишень. Если мы не знаем количества выстрелов, произведенных по мишени, скажем n, то нам приходится просто считать m значений x, наблюдаемых на мишени, полученными из распределения, сосредоточенного на интервале от 0 до R. Мы говорим тогда, что распределение x урезано справа в точке R. Аналогично, если мы определим в этом примере y как расстояние попадания от вертикальной линии, проходящей через центр мишени, то y может изме-

няться от $-R$ до R, и его распределение дважды урезано. Аналогично, мы можем иметь варианту, урезанную слева (например, если наблюдения, меньшие некоторого значения, не регистрируются). В общем случае варианта может быть многократно урезана одновременно в нескольких частях своей области изменения. Урезанная варианта ничем существенным не отличается от любой другой, но она рассматривается отдельно, так как ее распределение порождается исходной неурезанной случайной величиной, которая может иметь обычный вид. Так, в упражнении 17.26 мы рассматривали распределение Пуассона, урезанное слева, чтобы исключить нулевую частоту.

Тьюки (1949) и У. Смит (1957) показали, что урезание в фиксированных точках не меняет никаких свойств достаточности и полноты, которыми обладает статистика.

32.16 С другой стороны, рассмотрим снова наш пример с мишенью из **32.15**, но теперь предположим, что мы знаем, сколько выстрелов было произведено по мишени. Мы по-прежнему наблюдаем лишь m значений x, попавших от 0 до R включительно, но теперь мы знаем, что, помимо этого, существуют $n - m = r$ значений x и что эти значения превышают R. Другими словами, мы наблюдаем первые m порядковых статистик $x_{(1)}, \ldots, x_{(m)}$ в выборке объема n. Про выборку тогда говорят, что она *цензурирована справа* в точке R. (Цензурирование есть свойство выборки, тогда как урезание есть свойство распределения.) Аналогично, мы можем иметь *цензуризование слева* (например, при измерении реакции на некоторый возбудитель может быть необходима некоторая минимальная реакция для того, чтобы измерение было вообще возможно) и *двойное цензурирование*, где r_1 наименьших и r_2 наибольших значений в выборке объема n не наблюдаются и для оценки имеются только остальные $m = n - (r_1 + r_2)$.

Цензурированные выборки различаются следующим образом. В приведенных примерах цензурирование возникало из-за того, что значения варианты попадали за пределы некоторой наблюдаемой области; цензурирование имело место в некоторых фиксированных точках. Это называется цензурированием типа I. Говорят, что происходит цензурирование типа II, когда *фиксированная доля объема выборки* n цензурируется на нижнем и/или верхнем концах области значений x. На практике цензурирование типа II часто случается, когда наблюдаемая варианта x представляет собой период времени (например, период времени до разрушения изделия, проходящего испытание), а время на эксперимент ограничено. Тогда может быть решено прекращать эксперимент, когда получены первые m из n наблюдений. Отсюда видно, что цензурирование типа II обычно бывает справа.

С точки зрения теории основное различие между цензурированием типа I и типа II состоит в том, что в первом случае m (число наблюдений) есть случайная величина, тогда как в последнем случае оно фиксировано заранее. Соответственно теория цензурирования типа II проще.

Конечно, одностороннее урезание или цензурирование есть лишь частный случай двойного урезания или цензурирования, когда один из концов распределения не ограничивается, в то время как «обычная» ситуация является, так сказать, вдвойне крайним случаем, когда совсем нет никаких ограничений.

32.17 К настоящему времени имеется обширная литература по проблемам урезания и цензурирования. Подробное изложение этого предмета заняло бы слишком много места. Поэтому мы резюмируем результаты в пунктах **32.17—22** и предоставляем читателю, интересующемуся этим предметом, следовать литературным ссылкам. Мы классифицируем задачи оценивания по трем основным группам.

(A) *Оценки максимального правдоподобия.* Решение любой из задач может быть получено методом максимального правдоподобия; обычно уравнения правдоподобия решаются только итерационными методами. Например, если непрерывная варианта с функцией плотности $f(x|\theta)$ дважды урезана в известных точках a, b с $a < b$, то функция правдоподобия, если сделано n наблюдений, равна

$$L_T(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta) \Big/ \left\{ \int_a^b f(x|\theta)\,dx \right\}^n, \qquad (32.36)$$

причем знаменатель в (32.36) возникает из-за того, что урезанная варианта имеет ф. п.

$$f(x, \theta) \Big/ \int_a^b f(x|\theta)\,dx.$$

Максимум (32.36) может быть найден обычными методами.

Рассмотрим теперь ту же варианту, дважды цензурированную в фиксированных точках a, b, так что не наблюдаются r_1 наименьших и r_2 наибольших элементов выборки. Для этого цензурирования типа I функция правдоподобия равна

$$L_I(x|\theta) \propto$$

$$\propto \left\{ \int_{-\infty}^{a} f(x|\theta)\,dx \right\}^{r_1} \prod_{i=1}^{n-r_1-r_2} f(x_i|\theta) \left\{ \int_b^{\infty} f(x|\theta)\,dx \right\}^{r_2}, \qquad (32.37)$$

и r_1 и r_2, разумеется, являются случайными величинами.

С другой стороны, если имеет место цензурирование типа II с фиксированными r_1 и r_2, то функция правдоподобия равна

$$L_{\mathrm{II}}(x\,|\,\theta) \propto$$

$$\propto \left\{ \int_{-\infty}^{x_{(r_1+1)}} f(x\,|\,\theta)\,dx \right\}^{r_1} \prod_{i=r_1+1}^{n-r_2} f(x_{(i)}\,|\,\theta) \left\{ \int_{x_{(n-r_2)}}^{\infty} f(x\,|\,\theta)\,dx \right\}^{r_2}. \quad (32.38)$$

(32.37) и (32.38) имеют в точности одинаковую форму. Они отличаются тем, что пределы интегрирования в (32.38) случайны, а в (32.37) — нет и что r_1, r_2 случайны в (32.37), а в (32.38) — нет. Однако для данной совокупности наблюдений это формальное сходство позволяет применять одни и те же методы итерации для получения решений максимального правдоподобия. Кроме того, при $n \to \infty$ оба типа цензурирования асимптотически эквивалентны.

Б. Рао (1958а) показал при некоторых условиях регулярности, что цензурирование всегда приводит к потере эффективности оценивания, но что для урезания это может быть и не так. Это справедливо также, когда наблюдения группированы (Свами (1962а)). См. упражнение 32.26.

Гальперин (1952а) показал при условиях регулярности, аналогичных условиям в **18.16** и **18.26**, что МП-оценки при цензурировании типа II состоятельны, асимптотически нормальны и эффективны (см. упражнение 32.15).

Хартли (1958) дает общий метод итерационного решения уравнений правдоподобия для неполных данных (включая как урезание, так и цензурирование) из дискретных распределений.

(Б) *Несмещенные линейные оценки с минимальной дисперсией.* Второй подход состоит в том, чтобы искать линейную функцию от имеющихся в наличии порядковых статистик, которая была бы несмещенной оценкой интересующего нас параметра с минимальной дисперсией. Для этого мы пользуемся методом наименьших квадратов в применении к упорядоченным наблюдениям. Мы уже рассматривали теорию этого в случае, когда имеются все наблюдения, в **19.18—21**, и она может быть непосредственно применена к случаю урезания при условии, что вектор математических ожиданий и матрица рассеяния порядковых статистик вычисляются по урезанному распределению, а не по исходному распределению, над которым производилось урезание. Практическая трудность здесь состоит в том, что эта матрица рассеяния зависит от точек урезания a, b, так что несмещенная линейная функция с минимальной дисперсией будет меняться при изменении a и b. Работ в этой области мало или, может быть, нет совсем, по-видимому, из-за этой трудности.

Когда мы переходим к цензурированным выборкам, эта трудность сохраняется при цензурировании типа I, так как мы не знаем, сколько порядковых статистик окажется в пределах цензурирования (a, b). Таким образом, оценка должна быть определена отдельно для каждого значения r_1 и r_2, а ее математическое ожидание и дисперсия должны вычисляться по всем возможным r_1 и r_2 с соответствующей вероятностью для каждой комбинации. Снова мы не знаем случая, когда это было бы проделано. Однако при цензурировании типа II такой проблемы не возникает, так как r_1 и r_2 фиксированы заранее, и мы всегда знаем, какие $(n - r_1 - r_2)$ порядковых статистик имеются в наличии для целей оценивания. Зная их математические ожидания и матрицу рассеяния, мы можем непосредственно применить теорию НК из **19.18—21**. Кроме того, математические ожидания и матрицу рассеяния всех n порядковых статистик требуется вычислить лишь однажды для каждого n. Тогда для любых r_1, r_2 мы можем выбрать $(n - r_1 - r_2)$ математических ожиданий имеющихся наблюдений и подматрицу, являющуюся их матрицей рассеяния.

Чернов и др. (1967) получили общие формулы для того, чтобы линейные комбинации функций от порядковых статистик давали эффективные оценки параметров сдвига и масштаба в цензурированных или нецензурированных выборках.

Рядом авторов были предложены более простые процедуры, чтобы избежать вычислительных сложностей методов МП и НК. Наиболее общие результаты были получены Бломом (1958), который построил «почти» несмещенные «почти» эффективные линейные оценки, и Плэкеттом (1958), который показал, что МП-оценки для параметров сдвига и масштаба асимптотически линейны и что линейные несмещенные МД-оценки асимптотически нормально распределены и эффективны. Таким образом, по крайней мере асимптотически, эти два подхода смыкаются.

Гастверт (1966) исследовал взаимосвязь между линейными несмещенными МД-оценками и соответствующими асимптотически наиболее мощными критериями.

32.18 Мы теперь кратко изложим результаты, имеющиеся для каждого из основных распределений, которые изучались с точки зрения урезания и цензурирования; численные подробности слишком обширны, чтобы приводить их здесь.

Нормальное распределение. Свами (1962b) показал, что урезание всегда снижает эффективность, когда оцениваются и среднее и дисперсия, и (1963) то же происходит обычно, когда наблюдения группированы (см. Гранди (1952)). Для одинарного и двойного урезания оценивание методом МП рассматривалось Коэном (1950a, 1957), где были даны графики, облегчающие итерационное решение уравнений МП; Коэн и Вудворд (1953)

приводят таблицы, а Хальд (1949) и Гальперин (1952b) — графики для МП-оценивания в случае одинарного урезания. Итерационные процедуры МП для случая одинарного и двойного цензурирования типа II даются Хартером и Муром (1966a), где имеется также обзор более ранних работ. МП-оценки имеют тенденцию давать несколько бо́льшую точность, особенно когда цензурирование сильно несимметрично, чем несмещенные линейные МД-оценки, изучавшиеся Сарханом и Гринбергом (1956, 1958), чья книга (1962) содержит таблицы коэффициентов этих оценок для всех комбинаций чисел цензурирования r_1 и r_2, когда $n = 1(1)20$. Линеаризованные МП-оценки, предложенные Плэкеттом (1958), имеют эффективность, всегда не меньшую чем 99,98% при $n = 10$. Диксон (1957) показал, что при оценке среднего значения очень простая «обрубленная» *) оценка

$$t = \frac{1}{n-2} \sum_{i=2}^{n-1} x_{(i)}$$

имеет эффективность, всегда не меньшую, чем 99% при $n = 3(1)20$ и предположительно также при $n > 20$, тогда как среднее «двух наилучших» наблюдений (т. е. тех, среднее которых есть несмещенная оценка с минимальной дисперсией) имеет эффективность, медленно убывающую от 86,7% при $n = 5$ до своего асимптотического значения 81%. «Двумя наилучшими» наблюдениями являются приближенно $x_{(0,27n)}$ и $x_{(0,73n)}$ (см. упражнение 32.14). Аналогичные простые оценки стандартного отклонения σ даются статистиками

$$u = \sum_i \{x_{(n-i+1)} - x_{(i)}\}$$

с коэффициентами, обеспечивающими несмещенность, где суммирование включает 1, 2, 3 или 4 значения i. Наилучшая статистика такого типа при оценке σ всегда имеет эффективность не меньшую 96%.

Диксон (1960) показал, что если с каждой стороны цензурировано по i наблюдений, то «винзоризованная» оценка среднего

$$m_w = \frac{1}{n}\left[\sum_{s=i+2}^{n-i-1} x_{(s)} + (i+1)(x_{(i+1)} + x_{(n-i)})\right]$$

имеет эффективность по меньшей мере 99,9% по сравнению с несмещенной линейной оценкой с минимальной дисперсией, а

*) В оригинале «trimmed». (*Прим. ред.*)

при одностороннем цензурировании i наблюдений (скажем, справа) аналогичная оценка

$$m_a = \frac{1}{n+a-1}\left[\sum_{s=2}^{n-i-1} x_{(s)} + (i+1)\,x_{(n-i)} + a x_{(1)}\right],$$

где a выбрано так, чтобы сделать m_a несмещенной, эффективна по меньшей мере на 96%.

Некоторые общие результаты об эффективности обрубленных и винзоризованных оценок среднего для симметричных и симметричных унимодальных распределений имеются у Бикела (1965).

Уолш (1950а) показал, что оценка процентной точки нормального распределения с помощью подходящей порядковой статистики очень эффективна (хотя эта процедура оценки действительна на самом деле для любой непрерывной ф. р.) при одностороннем цензурировании типа II, когда цензурирована большая часть выборки.

Со (1959) показал, что в выборках с односторонним цензурированием типа II теоретическое среднее может быть оценено с асимптотической эффективностью, не меньшей 94%, посредством линейной комбинации наблюдений, ближайших к точке цензурирования (x_i), и простого среднего остальных наблюдений, а теоретическое стандартное отклонение оценивается с асимптотической эффективностью 100% с помощью суммы и суммы квадратов остальных наблюдений. Со приводит таблицы соответствующих весов для $n \leqslant 20$. Для цензурированных выборок типа I Со (1961) предложил простые линейные оценки высокой эффективности.

32.19 *Экспоненциальное распределение.* Распределение $f(x) = \exp\{-(x-\mu)/\sigma\}/\sigma$, $\mu \leqslant x \leqslant \infty$, было очень полно изучено с точки зрения урезания и цензурирования, и причина этого — в его важности для изучения долговечности некоторых изделий, в частности электрических и электронных элементов. Очень полная библиография этой области испытаний на долговечность дана Менденхоллом (1958) и дополнена Говиндараджулу (1964).

Оценивание σ (с известным μ) методом МП при одностороннем урезании или цензурировании типа I справа рассматривалось Димером и Вотоу (1955) (см. упражнение 32.16). Их результаты обобщались на цензурированные выборки из смесей нескольких экспоненциальных распределений Менденхоллом и Хейдером (1958). Для цензурирования типа II справа МП-оценка для σ дана Эпстейном и Собелом (1953), и оценка, являющаяся также несмещенной линейной МД-оценкой, — Сарханом (1955) (см. упражнения 32.17, 32.18).

Сархан и Гринберг (1957) приводят для объемов выборки до 10 таблицы коэффициентов несмещенных линейных оценок с минимальной дисперсией для σ отдельно и для (μ, σ) совместно для всех комбинаций цензурирования типа II на краях. Несмещенные оценки с минимальной дисперсией, основанные на одной или двух порядковых статистиках, даются Хартером (1961b), Сарханом и др. (1963) и Сиддики (1963); такие же оценки, основанные на трех, четырех и пяти порядковых статистиках, — Куллдорфом (1963), где приводится и некоторая общая теория. См. также Лорент (1963). Салех (1966) получил оценки, основанные на k порядковых статистиках.

32.20 *Распределение Пуассона.* Коэн (1954) дает МП-оценки и их асимптотические дисперсии для однократно и двукратно урезанных и (типа I) цензурированных выборок из распределения Пуассона и обсуждает ранние, менее общие, работы, относящиеся к этому распределению. Коэн (1960b) приводит таблицы и график для оценки методом МП, когда урезаются нулевые значения. Тейт и Гоэн (1958) получают несмещенную МД-оценку, когда урезание производится слева, и в частном случае, когда урезаны только нулевые значения, сравнивают ее с МП-оценкой (смещенной) и простой несмещенной оценкой, предложенной Плэкеттом (1953) (см. упражнения 32.20, 32.22—32.24).

Коэн (1960a) рассматривает МП-оценку пуассоновского параметра и параметра θ, когда доля θ наблюденных значений «1» ошибочно классифицирована как «0», и тот же автор (1960c) дает процедуру МП-оценивания, когда нулевые значения и (ошибочно) часть значений «1» урезаны.

32.21 *Другие распределения.* Для гамма-распределения с тремя параметрами

$$dF = \frac{1}{\Gamma(\beta)} \exp\{-\alpha(x-\mu)\}\{\alpha(x-\mu)\}^{\beta-1} d\{\alpha(x-\mu)\}$$

Чепмэн (1956) рассматривает урезание справа и предлагает упрощенные оценки для (α, β) при известном μ и для (α, β, μ) совместно. Коэн (1950b) рассматривал оценивание методом моментов в случае урезания. Радж (1953) и Ден Бродер (1955) рассматривали случаи цензурирования и урезания; последняя работа касается оценки параметра при урезании на каждом хвосте распределения. Уилк и др. (1962) дают МП-оценки для (α, β) при известном μ и для (α, β, μ) совместно в случае цензурирования типа II справа.

Сархан и Гринберг (1959) рассматривают несмещенные линейные МД-оценки для равномерных распределений (см. упражнение 32.25). Даунтон (1966) рассматривает распределение экстремальных значений, приведенное в упражнении 18.6. Говиндараджулу (1966) делает то же самое для симметрично цензу-

рированного двойного экспоненциального распределения. Финни (1949a), Райдер (1955), Сэмпфорд (1955), Уилкинсон (1961) и Шах (1961) рассматривают однократно урезанные биномиальное и отрицательное биномиальное распределения. Шах (1961) рассматривает дважды урезанное биномиальное распределение.

Хартер и Мур (1966b) рассматривают локальное МП-оценивание для цензурированных выборок из трехпараметрического логнормального распределения (с неизвестной начальной точкой); прямая МП-оценка бесконечна (см. упражнение 18.23).

Проверка гипотез в цензурированных выборках

32.22 В отличие от большого количества работ по оцениванию, рассмотренных в **32.17—21**, до сих пор очень мало сделано по проблемам проверки гипотез в случаях урезания и цензурирования. Эпстейн и Собел (1953) и Эпстейн (1954) обсуждают критерии для цензурированных экспоненциальных распределений; Ф. Дэвид и Джонсон (1954, 1956) дают различные простые критерии для цензурированных нормальных выборок, основанные на выборочных медианах и квантилях. Гальперин (1961a) дает простые доверительные интервалы для однократно цензурированных экспоненциальных и нормальных выборок. Гехан (1965a, b) (см. также Гальперин (1960)) распространил критерий Вилкоксона на цензурированные выборки. Гастверт (1965) получил асимптотически наиболее мощные ранговые критерии для цензурированных выборок.

Выпадающие наблюдения

32.23 В последних разделах этой главы мы кратко рассмотрим задачу, с которой в тот или иной момент встречается каждый практический статистик, а возможно, и большинство ученых-практиков. Задача состоит в том, чтобы решить, не получено ли одно или более из наблюдений из генеральной совокупности, отличной от той, которая порождает остальные наблюдения; она отличается от обычной задачи двух выборок тем, что мы не знаем заранее, какие наблюдения могут быть из другой совокупности, — если бы мы знали, то, разумеется, могли бы применить двухвыборочную технику, которая рассматривалась в предыдущих главах. Здесь же нам нужно выяснить, имело ли место вообще такое «засорение» выборки.

Постановка этой задачи обычно вызывается подозреваемой ошибкой инструмента или записи; ученый исследует свои данные обычным путем и подозревает, что некоторые (обычно только одно) из наблюдений имеют слишком экстремальные (высокие или низкие или и то и другое) значения, чтобы быть

в согласии с предположением, что они порождены тем же распределением. Теперь нужен какой-нибудь объективный метод, позволяющий решить, основательно ли это подозрение.

Поскольку подозрение ученого вызвано поведением на хвостах наблюдаемого им распределения, то «естественные» критерии, которые напрашиваются сами собой, основаны на поведении экстремальных порядковых статистик и, в частности, на их отклонении от некоторой характеристики положения для неподозреваемых наблюдений, либо (особенно когда подозреваются и «высокие» и «низкие» ошибки) сам выборочный размах может быть использован в качестве статистики критерия. Поэтому, например, Ирвин (1925) исследовал распределение $(x_{(p)} - x_{(p-1)})/\sigma$ в выборках из нормальной генеральной совокупности (см. также Пирсон (1926) и Силлито (1951)), а Стьюдент (1927) рекомендовал пользоваться размахом для проверки выпадающих наблюдений.

Начиная с этих ранних обсуждений задачи, в этом же направлении была проделана большая работ, причем практически всюду рассматривался только случай нормального распределения. Теперь ясно, что распределение экстремальных наблюдений чувствительно к форме исходного распределения (см. главу 14), так что эти процедуры вряд ли могут быть устойчивы к отклонениям от нормальности, но, вообще говоря, трудно сделать что-нибудь иное, чем принятие нормальности, так как то же возражение ввиду неустойчивости было бы справедливо и для любого другого исходного распределения.

32.24 Фергюсон (1961), следуя Диксону (1950), выдвигает две общие альтернативные гипотезы. Модель A (гипотеза отличия в сдвиге) состоит в том, что n независимых нормальных наблюдений x_i имеют общую дисперсию σ^2 и средние $\mathbf{M}(x_i) = \mu + \sigma \Delta a_{\nu_i}$, где a — известные константы, не все равные между собой, а $(\nu_1, \nu_2, \ldots, \nu_n)$ — неизвестная перестановка чисел от 1 до n. Модель B (гипотеза отличия в масштабе) состоит в том, что x_i независимы и нормальны с общим средним μ и дисперсиями $\mathbf{D}(x_i) = \sigma^2 \exp(\Delta a_{\nu_i})$. В каждой из моделей мы проверяем гипотезу H_0: $\Delta = 0$.

Рассматривая только критерии, инвариантные относительно изменений положения и масштаба, Фергюсон (1961) показал, что в модели A с неизвестным μ локально наиболее мощный критерий для H_0 против односторонней альтернативы H_1: $\Delta > 0$ основан на $\sqrt{b_1}$, выборочном коэффициенте асимметрии. Большие значения $\sqrt{b_1}$ являются критическими, если $k_3(a)$, третья k-статистика значений a, положительна, и малые значения $\sqrt{b_1}$ являются критическими, если $k_3(a) < 0$. Если (как и бывает в задачах о выпадающих наблюдениях) $n - l$ значений a равны

нулю и $l/n < 1/2$, то $k_3(a) > 0$, так что большие значения $\sqrt{b_1}$ всегда являются критическими, если выпадает менее половины наблюдений. Для двусторонней альтернативы H_1': $\Delta \neq 0$ критерий, основанный на выборочном коэффициенте эксцесса b_2, является локально наиболее мощным несмещенным, причем критическими являются большие или малые значения, когда $k_4(a) > 0$ или < 0 соответственно. Если $n - l$ значений a равны нулю, то $k_4(a) > 0$, когда $l/n \leqslant 0{,}21$, так что большие значения b_2 всегда являются критическими, если выпадают менее чем 21% наблюдений. Для модели B, где имеет смысл только односторонняя альтернатива H_1: $\Delta > 0$, поскольку малые изменения масштаба в задачах о выделяющихся наблюдениях бывают только вверх, локально наиболее мощный критерий основан на больших значениях b_2, каковы бы ни были a (так что здесь допускается любое число выпадающих наблюдений).

Однако «локально» наиболее мощный означает «вблизи $\Delta = 0$», так что требуется еще свидетельство эффективности этих критериев при больших сдвигах. Кроме того, у этих критериев имеется сильный конкурент с тех пор, как Полсон (1952) и Кудо (1956) показали, что для модели A с не более чем одним выпадающим наблюдением вероятность правильно отвергнуть выпадающее наблюдение максимизируется, если использовать в качестве критерия *стьюдентизованное экстремальное отклонение* (см. Томпсон (1935), Пирсон и Чандра Секар (1936))

$$t_n = \frac{x_{(n)} - \bar{x}}{s} \qquad \left(\text{или } t_1 = \frac{\bar{x} - x_{(1)}}{s}\right) \tag{32.39}$$

(где s^2 — оценка σ^2 по всем имеющимся наблюдениям) для односторонних альтернатив $\Delta > 0$ или $\Delta < 0$ соответственно. То же свойство справедливо для *стьюдентизованного максимального абсолютного отклонения*

$$t_{\max} = \max\{t_n, t_1\}, \tag{32.40}$$

используемого против двусторонних альтернатив $\Delta \neq 0$ в модели A, а также (Фергюсон (1961)) против $\Delta > 0$ в модели B.

Фергюсон (1961) провел выборочные эксперименты в модели A, основанные на 25 000 случайных нормальных отклонений, последовательно объединяемых в выборки объема $n = 5(5)25$, с одним выпадающим наблюдением, отличающимся от остальных на $\sigma(\sigma)15\sigma$. Среди односторонних критериев t_n оказался несколько лучше, чем $\sqrt{b_1}$, тогда как b_2 и $t_{\max}$ как двусторонние критерии отличались очень мало.

В *Biometrika Tables* даны 0,95- и 0,99-квантили для $\sqrt{b_1}$ при $n \geqslant 25$ и для b_2 при $n > 200$. Фергюсон (1961) оценивает их квантили с помощью выборочного эксперимента для $n = 5(5)25$. Куисенберри и Дэвид (1961) табулировали распределение t_1 (t_n) и $t_{\max}$.

32.25 Диксон (1950, 1951) рассматривал отношения вида $\dfrac{x_{(r)} - x_{(1)}}{x_{(n-s)} - x_{(1)}}$ в качестве критериев и проводил выборочные эксперименты для исследования их мощностей. Когда σ известна, он нашел, что нормированное экстремальное отклонение

$$u_n = \frac{x_{(n)} - \bar{x}}{\sigma} \qquad \left(\text{или } u_1 = \frac{\bar{x} - x_{(1)}}{\sigma}\right) \tag{32.41}$$

и нормированный размах

$$w = \frac{x_{(n)} - x_{(1)}}{\sigma} \tag{32.42}$$

имеют примерно одинаковую мощность. Дэвид и др. (1954) табулировал процентные точки стьюдентизованного размаха

$$q = \frac{x_{(n)} - x_{(1)}}{s}. \tag{32.43}$$

Распределение статистики (32.43) табулировано также в упрощенном случае, когда ее числитель и знаменатель получены из независимых выборок, в *Biometrika Tables* Пэчерсом (1959) и Хартером (1960) *). Распределение статистики (32.39) с тем же упрощением табулировано в *Biometrika Tables* Нэйром (1948, 1952), Дэвидом (1956), Пиллаи (1959) и Пиллаи и Тиензо (1959) (см. также Дэвид и Полсон (1965)). Аналогичным образом распределение статистики (32.40) табулировано Гальпериным и др. (1955). Энскомб (1960) исследовал влияние отбрасывания выпадающих наблюдений на последующую оценку, главным образом когда σ известно. Блисс и др. (1956) дали критерий размаха для браковки единственного выпадающего наблюдения среди k нормальных выборок объема n, а также таблицы к нему. Другие критерии рассматриваются Граббсом (1950) и Диксоном (1950, 1951, 1953а).

Диксон (1962, глава 10 Н книги Сархана и Гринберга (1962)) дает обширный обзор по этой теме, включая многие из упомянутых выше таблиц и некоторые другие.

Очевидный путь повышения устойчивости критериев и оценок к наличию выпадающих наблюдений состоит в том, чтобы основывать их на «центральной» части выборки, как, например, «обрубленные» и «винзоризованные» оценки в **32.18** и критерии в **32.22**.

Распределения, отличные от нормального

32.26 Один из немногих общих методов решения проблемы выпадающих наблюдений принадлежит Дарлингу (1952), который получил интегральную форму для х. ф. распределения статистики

$$z_n = \sum_{i=1}^{n} x_i / x_{(n)}, \tag{32.44}$$

где n наблюдений x_i суть независимые одинаково распределенные варианты с положительными значениями, распределение

*) Таблица этого распределения приведена в *Сборнике статистических таблиц*. Таблицы распределений многих статистик критериев для выпадающих наблюдений имеются в *Таблицах математической статистики*. (*Прим. ред.*)

которых полностью задано. В частных случаях эта х. ф. может быть обращена. Дарлинг далее подробно рассматривает случай варианты, имеющей распределение χ^2. Здесь мы рассмотрим только более простой случай равномерного распределения, в котором результат Дарлинга может быть получен непосредственно.

Предположим, что мы имеем наблюдения $x_1, x_2, \ldots, x_n$, равномерно распределенные на интервале $(0, \theta)$. Тогда мы знаем из **17.40**, что наибольшее наблюдение $x_{(n)}$ есть достаточная статистика для θ, а из **23.12** — что $x_{(n)}$ есть полная достаточная статистика. Следовательно, согласно результату упражнения 23.7 любая статистика, распределение которой не зависит от θ, будет независима от $x_{(n)}$. Ясно, далее, что z_n, определенная в (32.44), имеет нулевую степень по θ. Таким образом, z_n не зависит от $x_{(n)}$ и условное распределение z_n при фиксированном $x_{(n)}$ — такое же, что и безусловное (маргинальное) распределение. Но при фиксированном $x_{(n)}$ любое x_i $(x_i < x_{(n)})$ распределено равномерно на интервале $(0, x_{(n)})$. Таким образом, $x_i/x_{(n)}$ при фиксированном $x_{(n)}$ имеет равномерное распределение на интервале (0, 1), и мы видим из (32.44), что z_n распределено в точности, как сумма $n-1$ независимых случайных величин, распределенных равномерно на (0, 1), плюс постоянная $1\ (= x_{(n)}/x_{(n)})$. Поскольку мы видели в примере 11.9, что распределение суммы n независимых равномерных случайных величин стремится к нормальному (и ее распределение в самом деле близко к нормальному даже при $n = 3$), z_n распределена асимптотически нормально со средним и дисперсией, даваемыми точными выражениями

$$\mathbf{M}(z_n) = (n-1)\frac{1}{2} + 1 = \frac{1}{2}(n+1), \qquad \mathbf{D}(z_n) = \frac{1}{12}(n-1). \quad (32.45)$$

Малые значения z_n (отвечающие большим значениям $x_{(n)}$) образуют критическую область для проверки гипотезы, что все наблюдения распределены одинаково, против альтернативы, что наибольшее из них происходит из «выпадающего» распределения.

32.27 Результат Дарлинга можно использовать для выявления выпадающего наблюдения при любом полностью заданном исходном распределении, сделав сначала вероятностное интегральное преобразование наблюдений (см. **30.36**), приводящее задачу к равномерному распределению на (0, 1). Наименьшее наблюдение $x_{(1)}$ можно проверять аналогичным образом, беря дополнения до 1 этих равномерных величин и проверяя $x_{(n)}$, как раньше.

Басу (1965) рассматривает выпадающие наблюдения для экспоненциального распределения.

32.28 Мы должны теперь указать на возможность использования свободных от распределения методов для решения задачи о выпадающих наблюдениях без конкретных предположений о распределениях. Ясно, что, если экстремальные наблюдения находятся под подозрением, это автоматически сводит на нет любую попытку использовать для этой задачи обычный двухвыборочный критерий, основанный на порядке рангов, какие рассматривались в главе 31. Однако если мы готовы предположить симметрию (непрерывного) исходного распределения, то мы в состоянии кое-что сделать, так как мы можем тогда сравнивать поведение наблюдений на подозреваемом «хвосте» наблюдаемого распределения с поведением на другом «хвосте», который, как мы предполагаем, ведет себя хорошо. Таким образом, при больших n мы можем рассматривать абсолютные отклонения от выборочного среднего (или медианы) k наибольших и k наименьших наблюдений, ранжировать эти $2k$ значений и применить критерий симметрии, чтобы решить, можно ли их считать однородными. Критерий будет приближенным, так как центр симметрии неизвестен и мы оцениваем его посредством выборочного среднего или медианы, но в остальном это просто применение критерия симметрии (см. **31.78—79**) к хвостам распределения. Если n достаточно велико, а k настолько велико, чтобы иметь возможность разумно выбрать размер критерия α, то эта процедура должна быть достаточно чувствительной для практических целей.

По существу аналогичные, но более сложные критерии того, следует ли считать группу из четырех или более наблюдений выпадающей, были предложены Уолшем (1950b).

32.29 Наконец, заметим, что свободные от распределения методы проверки гипотез и оценивания подвержены меньшему влиянию выпадающих наблюдений, чем методы, основанные на предположениях о распределении, так как они используют порядковые, а не метрические свойства наблюдений, как мы видели в главе 31.

УПРАЖНЕНИЯ

32.1 Для функции плотности $f(x)$, принимающей наибольшее значение в медиане M, показать, что АОЭ критерия знаков по сравнению с t-критерием Стьюдента, приведенная в (32.12), никогда не меньше чем 1/3 и это значение достигается, когда $f(x)$ прямоугольна.

(Ходжес и Леман, 1956.)

32.2 Показать, что если n независимых наблюдений получены из одного и того же непрерывного распределения $F(x)$, то любая симметричная функция от них стохастически независима с любой функцией от их порядка рангов. Как следствие этого показать, что комбинация критерия знаков и критерия ранговой корреляции может быть использована для проверки гипотезы, что $F(x)$ имеет медиану M_0, против альтернативы, что либо наблюдения

одинаково распределены с медианой $\neq M_0$, либо медиана имеет тренд вверх (или вниз) для каждого последующего наблюдения.

(Сэвидж, 1957.)

32.3 Получить результат (32.12) для АОЭ критерия знаков в проверке симметрии из эффективности выборочной медианы по отношению к выборочному среднему при оценке центра симметричного распределения (см. **25.13**).

32.4 Показать, пользуясь (32.23), что при построении доверительных интервалов для медианы непрерывного распределения с помощью симметрично расположенных порядковых статистик $x_{(r)}$ и $x_{(n-r+1)}$ при $n = 30$ указанные ниже значения r дают указанные коэффициенты доверия.

r	$1-\alpha$	r	$1-\alpha$
8	0,995	12	0,80
9	0,98	13	0,64
10	0,96	14	0,42
11	0,90	15	0,14

32.5 Показать, что при проверке гипотезы H_0: $\theta = \theta_0$ против альтернативы H_1: $\theta = \theta_1 > \theta_0$ по выборке из n наблюдений из распределения

$$dF = \frac{1}{2}\exp\{-|x-\theta|\}\,dx, \qquad -\infty \leqslant x \leqslant \infty,$$

односторонний критерий знаков является асимптотически наиболее мощным.

(См. Леман, 1959.)

32.6 На основе примера 32.1 показать, что размах выборки объема $n = 100$ из непрерывного распределения имеет вероятность, превышающую 0,95, покрыть не меньше чем 95% исходной ф. р., но что если мы хотим покрыть не меньше чем 99% с вероятностью 0,95, то n должно быть приблизительно 475 или более.

32.7 На основе примера 32.1 показать, что если мы хотим найти свободный от распределения толерантный интервал для доли γ непрерывной ф. р. с вероятностью β, близкой к 1, то малое увеличение β с β_1 до β_2 требует увеличения объема выборки приближенно от n_1 до $n_2 = n_1\left(\frac{1-\beta_2}{1-\beta_1}\right)$.

32.8 Пусть $F(x, \theta)$ — любая непрерывная ф. р., $F(a|\theta) = 0$, $F(b|\theta) = 1$ и значения $\lambda_0, \lambda_1, \ldots, \lambda_{n+1}$ определены соотношением $F(\lambda_i|\theta_0) = \frac{i}{n+1}$, где θ_0 — истинное значение θ. Рассмотрим выборку из n наблюдений и разобьем выборочное пространство на $(n+1)^n$ частей, имеющих вероятности P_r, $r = 1, 2, \ldots, (n+1)^n$, плоскостями $x_i = \lambda_j$, $i = 1, 2, \ldots, n$; $j = 0, 1, \ldots, n+1$. Показать, что если t — любая асимптотически несмещенная оценка θ, а t_r — ее значение в некоторой произвольно выбранной точке r-й из $(n+1)^n$ частей выборочного пространства, то при $n \to \infty$

$$\mathrm{D}t \sim \sum_r (t_r - \theta)^2 P_r \geqslant \left[n^2 \sum_{i=0}^{n} \left\{\left(\frac{\partial F}{\partial \theta}\right)_{\lambda_{i+1}} - \left(\frac{\partial F}{\partial \theta}\right)_{\lambda_i}\right\}^2\right]^{-1}.$$

(Блом, 1958.)

32.9 Показать, что если $\partial F/\partial\theta$ — непрерывная функция от x в области изменения x, стремящаяся к нулю на ее краях, и интеграл $\mathbf{M}\left(\frac{\partial \log f}{\partial \theta}\right)^2 > 0$

существует, где f — ф. п. $\dfrac{\partial F(x\mid\theta)}{\partial x}$, то асимптотическая нижняя граница для дисперсии несмещенной оценки, даваемая в упражнении 32.8, переходит в МГД (17.24).

32.10 Показать, что при оценке среднего значения равномерного распределения с известным размахом результат упражнения 32.8 дает для дисперсии асимптотическую границу $(2n^2)^{-1}$ и она достигается на выборочной середине размаха $t=\frac{1}{2}(x_{(1)}+x_{(n)})$.

32.11 Показать, что если $\partial F/\partial\theta$, рассматриваемая как функция от x, имеет счетное число разрывов и $\mathbf{M}\left(\dfrac{\partial \log f}{\partial\theta}\right)^2$ существует, то асимптотическая граница дисперсии из упражнения 32.8 имеет порядок n^{-2}.

32.12 Показать, что для выборок объема n из логистического распределения

$$dF=e^{-x}\,dx/(1+e^{-x})^2, \qquad -\infty\leqslant x\leqslant\infty,$$

х. ф. r-й порядковой статистики $x_{(r)}$ равна

$$\varphi_r(t)=\frac{\Gamma(r+it)\,\Gamma(n-r+1-it)}{\Gamma(r)\,\Gamma(n-r+1)},$$

так что, из (16.29), $\frac{1}{2}\left(x_{(r)}-\log\dfrac{n-r+1}{r}\right)$ имеет z-распределение Фишера (16.26) с $\nu_1=2r$, $\nu_2=2(n-r+1)$, а следовательно, семиинварианты $x_{(r)}$ равны при $r>\frac{1}{2}(n+1)$

$$\varkappa_1=\sum_{s=n-r+1}^{r-1}\frac{1}{s}, \qquad \varkappa_2=\frac{\pi^2}{3}-\left(\sum_{s=1}^{r-1}\frac{1}{s^2}+\sum_{s=1}^{n-r}\frac{1}{s^2}\right),$$

$$\varkappa_3=2\sum_{s=n-r+1}^{r-1}\frac{1}{s^3}, \qquad \varkappa_4=\frac{2\pi^4}{15}-6\left(\sum_{s=1}^{r-1}\frac{1}{s^4}+\sum_{s=1}^{n-r}\frac{1}{s^4}\right).$$

(Плэкетт, 1958; А. Бирнбаум и Дадмэн (1963) — таблица средних значений и стандартных отклонений для $n=1(1)10(5)20,\ 50,\ 100$; Гупта и Шах (1965) — таблица первых четырех моментов с набором квантилей для распределений всех порядковых статистик, когда $n=1(1)10$, и квантилей только для крайних и центральных порядковых статистик, когда $n=11(1)25$. Тартер и Кларк (1965) дают формулы для ковариаций логистических порядковых статистик, табулированных для $n\leqslant 10$ Б. Шахом (1966).)

32.13 Непрерывное распределение имеет ф. р. $F(x)$, и ξ определяется соотношением $F(\xi)=p$. Показать, разлагая x в ряд Тейлора относительно

значения $\mathbf{M}\{x_{(r)}\}$, что в выборках объема n, когда $n \to \infty$, при $r = [np]$

$$\mathbf{M}(x_{(r)}) = p + O(n^{-1})$$

и

$$F[\mathbf{M}\{x_{(r+1)}\}] - F[\mathbf{M}\{x_{(r)}\}] = \frac{1}{n} + O(n^{-2}).$$

(Плэкетт, 1958.)

32.14 Показать, что в выборках из нормальной совокупности, если оценивать μ статистикой

$$t = \frac{1}{2}(x_{(i)} + x_{(n-i+1)}),$$

мы получим асимптотически минимальную дисперсию, выбирая $i = 0{,}2703n$. Аналогично, если мы оцениваем σ статистикой

$$s = \frac{1}{k}(x_{(n-i+1)} - x_{(i)}),$$

то мы получим минимальную дисперсию, когда $i = 0{,}0692n$.

(Бенсон, 1949; эти результаты были впервые даны К. Пирсоном в 1920 г.; см. также П. Мур (1956).)

32.15 Показать (см. (32.37), (32.38)), что влияние цензурирования аналогично приданию дискретной вероятности исходному распределению $f(x|\theta)$ в каждой области, где производится цензурирование. Получить отсюда, что, например, при цензурировании типа I выше фиксированной точки x_0 асимптотическая дисперсия МП-оценки для θ равна

$$\mathbf{D}\hat{\theta} \sim 1 \Big/ \left[n \left\{ \int_{-\infty}^{x_0} \left(\frac{\partial \log f}{\partial \theta} \right)^2 f\, dx + P \left(\frac{\partial \log P}{\partial \theta} \right)^2 \right\} \right],$$

где

$$P = \int_{x_0}^{\infty} f\, dx.$$

(См. Гальперин, 1952а.)

32.16 Показать, что для выборки из распределения

$$dF = \theta e^{-\theta x}, \qquad 0 \leqslant x \leqslant \infty; \quad \theta > 0,$$

содержащей $n + m$ наблюдений, из которых n измерены от 0 до фиксированного значения x_0, а m остальных, имеющие значения, большие x_0, подвергнуты цензурированию типа I, МП-оценка для θ равна

$$\hat{\theta} = n \Big/ \left\{ \sum_{i=1}^{n} x_{(i)} + m x_0 \right\}$$

и

$$\mathbf{D}\hat{\theta} \sim \frac{\theta^2}{n} \Big/ \{1 - \exp(-\theta x_0)\},$$

так что асимптотическая дисперсия $\hat{\theta}$ является монотонно убывающей функцией от x_0.

(Димер и Вотоу, 1955.)

32.17 В упражнении 32.16 показать, что если производится цензурирование типа II с фиксированным числом m выборочных порядковых статистик, цензурированных справа, то МП-оценкой для $\lambda = 1/\theta$ будет

$$\hat{\lambda} = \frac{1}{n}\left\{\sum_{i=1}^{n} x_{(i)} + m x_{(n)}\right\}$$

и что ее асимптотическая дисперсия равна

$$\mathrm{D}\hat{\lambda} \sim \lambda^2/n.$$

(Эпстейн и Собел, 1953.)

32.18 В упражнении 32.17 показать, что выражение для дисперсии $\hat{\lambda}$ является точным при любом n и что $\hat{\lambda}$ есть несмещенная линейная МД-оценка для λ.

(Сархан, 1955.)

32.19 Показать, что если распределение Пуассона с параметром θ урезано справа, то не существует несмещенной оценки для θ.

(Тейт и Гоэн, 1958.)

32.20 Показать, что если урезание удаляет значения $\leqslant a$, то для параметра θ распределения Пуассона имеется несмещенная оценка

$$t = \frac{1}{n}\sum_{r=a+2}^{\infty} r n_r,$$

где n_r — наблюденная частота значения r. Показать, что

$$\mathrm{D}t = \frac{\theta}{n}\left\{1 + \frac{\theta^{a+1}}{a!\left(e^{\theta} - \sum_{r=0}^{a}\frac{\theta^r}{r!}\right)}\right\}$$

и что несмещенной оценкой этой дисперсии является

$$\widehat{\mathrm{D}}t = \frac{1}{n}\left\{t + \frac{(a+1)(a+2)\,n_{a+2}}{n}\right\}.$$

(См. Субрахманиам (1965), где показано, что t имеет эффективность $\geqslant 90\%$ по сравнению с МП-оценкой, когда $\theta \leqslant 1$. Если $a = 0$, то t есть оценка Плэкетта (1953) и имеет эффективность $\geqslant 95\%$ при всех θ; см. также Дж. Рой и Митра (1957).)

32.21 Для выборки $x_1, \ldots, x_n$ показать, что

$$\sum_{r=1}^{n} x_r^2 \equiv n\bar{x}^2 + \sum_{r=1}^{n-1}\frac{z_r^2}{r(r+1)},$$

где

$$z_r = r x_{(r+1)} - \sum_{s=1}^{r} x_{(s)}, \qquad r = 1, 2, \ldots, n-1.$$

Вывести отсюда, применяя преобразование Хельмерта примера 11.3 к n порядковым статистикам в форме

$$y_r = z_r/\{r\,(r+1)\}^{1/2}, \qquad r = 1,\ 2,\ \ldots,\ n-1;\quad y_n = n^{1/2}\bar{x},$$

что в выборках из стандартного нормального распределения совместная функция плотности z_r равна

$$n^{1/2}\,(2\pi)^{-(n-1)/2} \exp\left\{-\frac{1}{2}\sum_{r=1}^{n-1} z_r^2/[r\,(r+1)]\right\},$$

$$0 \leqslant z_1 \leqslant z_2 \leqslant \ldots \leqslant z_{n-1} = n\,(x_{(n)} - \bar{x}).$$

Определяя функции

$$G_r(x) = \int_0^x \exp\left\{-\frac{1}{2}\,\frac{t^2}{r\,(r+1)}\right\} G_{r-1}(t)\,dt, \quad G_0(x) \equiv 1,$$

показать, что функция распределения $P_n(u)$ величины $u = x_n - \bar{x}$ удовлетворяет соотношению

$$P_n(u) = n^{1/2}\,(2\pi)^{-(n-1)/2}\,G_{n-1}(nu).$$

(Нэйр, 1948; Мак Кей (1935) получил эквивалентный результат.)

32.22 В упражнении 32.20 показать, что МП-оценка является корнем уравнения

$$\theta/(1 - e^{-\theta}) = \bar{x},$$

где $\bar{x}$ — выборочное среднее, и

$$\mathrm{D}\hat{\theta} \sim \frac{\theta}{n}\,(1 - e^{-\theta})^2/\{1 - (\theta+1)\,e^{-\theta}\}.$$

Вывести отсюда, что

$$\lim_{\theta\to 0} n\mathrm{D}\hat{\theta} = 2, \qquad \lim_{\theta\to\infty} n\mathrm{D}\hat{\theta} = 1$$

и что $n\mathrm{D}\hat{\theta}$ никогда не выходит за эти пределы.

(См. Коэн, 1960b.)

32.23 Показать, что МП-оценка в упражнении 32.22 может быть представлена в виде

$$\hat{\theta} = \bar{x} + \sum_{r=1}^{\infty} \frac{(-1)^r\,\bar{x}^r}{r!}\,\frac{d^{r-1}}{d\bar{x}^{r-1}}\,e^{-r\bar{x}} = \bar{x} - \sum_{r=1}^{\infty} \frac{r^{r-1}}{r!}\,(\bar{x}e^{-\bar{x}})^r.$$

(Ирвин, 1959.)

32.24 В упражнении 32.20 показать, что оценка числа наблюдений, включая отсутствующие значения, дается отношением $\left(\sum x\right)^2/\sum x\,(x-1)$, где суммирования производятся от 1 до ∞. Показать отсюда, как можно оценивать θ итеративно.

(Ирвин, 1959; метод принадлежит Мак Кендрику.)

32.25 В случае цензурированных типа II выборок из равномерного распределения $dF = dx/\sigma$, $\mu - \frac{1}{2}\sigma \leqslant x \leqslant \mu + \frac{1}{2}\sigma$, показать с помощью (32.38), что крайние наблюденные значения $x_{(r_1+1)}$, $x_{(n-r_2)}$ образуют пару достаточных статистик для μ и σ. Вывести отсюда, что несмещенными линейными МД-оценками для этих параметров являются

$$m = \frac{(n - 2r_2 - 1)\, x_{(r_1+1)} + (n - 2r_1 - 1)\, x_{(n-r_2)}}{2(n - r_1 - r_2 - 1)},$$

$$s = \frac{n+1}{n - r_1 - r_2 - 1}\left(x_{(n-r_2)} - x_{(r_1+1)}\right)$$

и их дисперсии равны

$$\mathsf{D}m = \frac{\sigma^2\{(r_1+1)(n - 2r_2 - 1) + (n - 2r_1 - 1)(r_2 + 1)\}}{4(n+1)(n+2)(n - r_1 - r_2 - 1)},$$

$$\mathsf{D}s = \frac{\sigma^2(r_1 + r_2 + 2)}{(n+2)(n - r_1 - r_2 - 1)},$$

что является обобщением примера 19.10.

(Сархан и Гринберг, 1959.)

32.26 Ссылаясь на **18.16**, показать, что если исходное распределение $f(x|\theta)$ урезано в точках a снизу и b сверху, то обратная величина асимптотической дисперсии $\hat{\theta}$ равна

$$R^2_{a,b}(\theta) = n\left[\left\{\int_a^b \frac{1}{f}\left(\frac{\partial f}{\partial\theta}\right)^2 dx \Big/ \int_a^b f\,dx\right\} - \left\{\left(\int_a^b \frac{\partial f}{\partial\theta}\,dx\right)^2 \Big/ \left(\int_a^b f\,dx\right)^2\right\}\right]$$

и стремится к предыдущему определению (18.30) при $a \to -\infty$, $b \to +\infty$. Показать, что $R^2_{-\infty,\infty}(\theta) - R^2_{a,b}(\theta)$ не обязательно положительна, но что всегда $R^2_{-\infty,\infty}(\theta) - \left(\int_a^b f\,dx\right) R^2_{a,b}(\theta) > 0$, так что мы можем ожидать некоторого уменьшения асимптотической дисперсии МП-оценки, если только урезание не слишком сурово.

(См. Б. Рао, 1958а.)

ГЛАВА 33

КАТЕГОРИЗОВАННЫЕ ДАННЫЕ

33.1 Бóльшая часть этой книги была посвящена анализу наблюдений, представленных в виде числовых значений. Однако в главе 31 для проверки гипотез были использованы также ранговые статистики. Последние могут быть получены как из числовых значений наблюдений, так и из рангов, если измерения получить трудно или даже невозможно. В настоящей главе рассматриваются категоризованные данные, т. е. данные, представленные в виде частот наблюдений, попавших в некоторые категории или классы. Мы, правда, уже имели дело с задачами группировки значений варианты в классы, например, в связи с вычислением моментов. Однако эта группировка производилась лишь для удобства вычислений, и, во всяком случае, мы в основном имели дело с одномерными задачами. Сейчас же мы будем специально заниматься вопросами, касающимися статистических связей (зависимости или взаимозависимости) между двумя или более «переменными», выраженными в категоризованной форме. Мы поставили слово «переменные» в кавычки, поскольку оно здесь понимается в самом общем смысле.

Категоризованная «переменная» может представлять собой просто подходящую классификацию числовой переменной по группам. С такой ситуацией мы уже знакомы. С другой стороны, такая переменная может вообще не выражаться в терминах какой-либо исходной числовой переменной. Например, можно классифицировать людей (а) по их росту, (б) по цвету волос, (в) по любимым киноактерам. В этом случае (а) будет категоризацией числовой переменной, чего нельзя сказать о (б) и (в). В свою очередь, (б) отличается от (в) тем, что цвет волос может быть в принципе представлен на шкале, упорядоченной в соответствии с пигментацией (от светлого к темному), тогда как в случае (в) это невозможно. Конечно, можно предложить различные виды классификации киноактеров, однако, по существу, киноактеры никак не упорядочены. Классификация, или категоризация, (б) является упорядоченной, а (в) — неупорядоченной. В качестве крайнего случая неупорядоченной классификации можно рассмотреть классификацию, заключающуюся

просто в различении нескольких выборок (которые мы хотим сравнивать по некоторой другой переменной).

33.2 Следует иметь в виду, что иногда две изучаемые переменные могут быть одной и той же переменной, наблюдаемой в двух разных ситуациях (например, до и после некоторого события) или в двух связанных между собой выборках (например, отец и сын, муж и жена и т. д.). В таких случаях мы будем говорить об *однородных категоризациях*. Разумеется, однородные категоризации могут принадлежать к любому из типов (а), (б) или (в) пункта **33.1**.

Связь в таблицах 2×2

33.3 Исторически очень большая часть литературы по категоризованным переменным была посвящена задачам проверки наличия и измерения взаимозависимости между двумя такими переменными. Мы полностью оставляем в стороне задачу оценивания взаимозависимости в случае, когда переменные — случайные величины, вид распределений которых известен или предполагается известным. В случае двумерного нормального распределения эта задача была рассмотрена в **26.27—33**. Другими словами, мы ограничиваемся непараметрическими задачами.

33.4 Рассмотрим вначале генеральную совокупность, в которой классификация произведена на основании наличия или отсутствия некоторого признака A. Простейшая задача о взаимозависимости возникает, когда имеется два признака A и B. Если α обозначает отсутствие A и β — отсутствие B, то количества попаданий в четыре возможные подгруппы могут быть в очевидных обозначениях представлены таблицей

	B	не-B	Сумма
A	(AB)	$(A\beta)$	(A)
не-A	(αB)	$(\alpha\beta)$	(α)
Сумма	(B)	(β)	n

(33.1)

Эта *таблица* 2×2 (иногда называемая *четырехклеточной таблицей*) часто будет записываться в уже встречавшейся нам форме (26.58):

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	n

(33.2)

Если между A и B не существует никакой связи, т. е. если обладание признаком A не связано с обладанием признаком B, то доля индивидов с признаком A среди индивидов, обладающих признаком B, должна быть равна доле индивидов с признаком A среди индивидов, не обладающих признаком B. Таким образом, по определению признаки независимы *) в данной совокупности из n наблюдений, если

$$\frac{a}{a+c}=\frac{b}{b+d}=\frac{a+b}{n}. \qquad (33.3)$$

Из (33.3), кроме того, получаем

$$\left.\begin{aligned}\frac{c}{a+c}&=\frac{d}{b+d}=\frac{c+d}{n},\\ \frac{a}{a+b}&=\frac{c}{c+d}=\frac{a+c}{n},\\ \frac{b}{a+b}&=\frac{d}{c+d}=\frac{b+d}{n}.\end{aligned}\right\} \qquad (33.4)$$

Соотношение (33.3) может быть также переписано в виде равенства

$$a=\frac{(a+b)(a+c)}{n}. \qquad (33.5)$$

Если теперь для какой-либо таблицы выполнено неравенство

$$a>\frac{(a+b)(a+c)}{n}, \qquad (33.6)$$

означающее, что доля A среди B больше, чем среди не-B, то A и B называются *положительно связанными* или просто *связанными*. В случае выполнения *противоположного* неравенства

$$a<\frac{(a+b)(a+c)}{n} \qquad (33.7)$$

мы будем говорить, что A и B *отрицательно связаны*.

*) Возможно, для выражения соотношения (33.3) лучше было бы использовать вместо слова «независимый» какое-нибудь более нейтральное слово, скажем «несвязанный», поскольку (33.3) не влечет стохастической независимости числовых переменных, порождающих таблицу 2×2 (хотя само это соотношение следует из стохастической независимости). Это различие аналогично различию между отсутствием корреляции и независимостью (см. **26.10**). Однако использование другого термина противоречило бы исторической традиции; кроме того, могла бы возникнуть путаница между понятием «несвязанный» («unassociated») и понятием «отрицательно связанный» («dissociated»), которое определяется далее в **33.4**. Мы поэтому будем продолжать использовать термин «независимый» в применении к категоризованным переменным, как указывающий на «отсутствие связи».

Пример 33.1

В следующей таблице (Гринвуд и Юл, 1915, *Proc. Roy. Soc. Medicine*, **8**, 113) приведены 818 случаев, классифицированных по двум признакам: наличию прививки против холеры (признак A) и отсутствию заболевания (признак B).

	Незаболевшие	Заболевшие	Сумма
Привитые	276	3	279
Непривитые	473	66	539
Сумма	749	69	818

Если бы признаки были независимы, то частота класса AB (включающего имеющих прививку и незаболевших) была бы равна $\frac{279 \times 749}{818} = 225$. Наблюденная частота больше этой величины. Следовательно, вакцинация положительно связана с отсутствием заболевания.

Меры связи

33.5 Если ставить вопрос о выражении степени связи двух признаков одним коэффициентом, то естественно потребовать, чтобы были известны границы изменения этого коэффициента и чтобы он принимал среднее или нижнее значение интервала изменения в случае, когда признаки не связаны («независимы»). Изменением положения и масштаба любой такой коэффициент можно заставить изменяться в интервале $(-1, +1)$, тогда случай независимости будет соответствовать нулевому значению коэффициента. Это удобно тем, что свойства определяемого коэффициента оказываются согласованными со свойствами обычного коэффициента корреляции (см. **26.9**). В некоторых случаях, однако, более предпочтительным является интервал $(0, 1)$; тогда нулевое значение соответствует случаю независимости.

Другим желательным свойством меры связи было бы ее возрастание при переходе от отрицательной связи к положительной. Рассмотрим разность между наблюденной частотой и частотой, полученной в предположении «независимости», для клетки (AB):

$$D = a - \frac{(a+b)(a+c)}{n} = \frac{ad - bc}{n}. \qquad (33.8)$$

Для постоянных маргинальных частот, очевидно, разность между наблюденной и «независимой» частотами в любой клетке

равна $\pm D$, тем самым D однозначно определяет отклонение от независимости. Поэтому естественно потребовать, чтобы наш коэффициент возрастал с возрастанием D.

33.6 Следуя Юлу (1900, 1912), мы определим коэффициент связи Q равенством

$$Q = \frac{ad - bc}{ad + bc} = \frac{nD}{ad + bc}. \tag{33.9}$$

Этот коэффициент равен нулю, если признаки независимы, т. е. если $D = 0$, и может принимать значение $+1$, только когда $bc = 0$, т. е. в случае *полной связанности* (либо все A одновременно являются B, либо все B одновременно являются A), а значение -1, только когда $ad = 0$, т. е. в случае полной отрицательной связанности. Кроме того, Q возрастает с возрастанием D. Действительно, вводя обозначение $e = bc/(ad)$, получаем

$$Q = (1 - e)/(1 + e) = 2/(1 + e) - 1,$$

откуда

$$\frac{dQ}{de} = -\frac{2}{(1 + e)^2} < 0.$$

Поскольку произвольная dD/de также отрицательна, то dQ/dD положительна. Юл предложил, кроме того, так называемый *коэффициент коллигации*

$$Y = \frac{1 - \left(\frac{bc}{ad}\right)^{1/2}}{1 + \left(\frac{bc}{ad}\right)^{1/2}} = \frac{(ad)^{1/2} - (bc)^{1/2}}{(ad)^{1/2} + (bc)^{1/2}}. \tag{33.10}$$

Но нетрудно показать, что

$$Q = \frac{2Y}{1 + Y^2}, \tag{33.11}$$

так что мы, по-видимому, ничего не выиграем, используя Y. Легко видеть, что коэффициент коллигации удовлетворяет всем нашим условиям.

Третий коэффициент, к которому мы вернемся позднее в **33.17**, определяется формулой

$$V = \frac{(ad - bc)}{\{(a + b)(a + c)(b + d)(c + d)\}^{1/2}}. \tag{33.12}$$

Очевидно, что он равен нулю, когда $D = 0$, и возрастает с возрастанием D. Если $V^2 = 1$, то мы получаем

$$(a + b)(a + c)(b + d)(c + d) = (ad - bc)^2,$$

что дает

$$4abcd + a^2(bc + bd + cd) + b^2(ac + ad + cd) + c^2(ab + ad + bd) + d^2(ac + ab + bc) = 0.$$

Так как ни одна из частот не может быть отрицательной, то это уравнение имеет решение только в случае, когда по крайней мере два из чисел a, b, c, d равны нулю. Случай, когда равны нулю частоты из одной строки или из одного столбца, является вырожденным. Остается, следовательно, рассмотреть случай $a = 0$, $d = 0$ и случай $b = 0$, $c = 0$. В первом из них $V = -1$, во втором $V = +1$. (Эти два числа служат для V границами.)

33.7 Можно заметить, что, в то время как $|V| = 1$ только при обращении в нуль двух частот в таблице 2×2, $|Q|$ и $|Y|$ принимают значение единица при обращении в нуль одной частоты. Здесь возникает вопрос об определении *полной* связи. Будем говорить, что связь между A и B полная, если все A являются одновременно B, несмотря на то что не все B являются одновременно A. Если все немые люди глухи, то это означает, что существует полная связь между немотой и глухотой, хотя есть много глухих людей, которые не являются немыми*). Коэффициент V равен 1 только тогда, когда все A являются одновременно B и все B являются одновременно A, — условие, которое можно, если угодно, рассматривать как определение *абсолютной* связи.

Следует обратить внимание на то, что статистическая связь отличается от связи в обычном смысле. В повседневной речи мы говорим, что A и B связаны, если они достаточно часто встречаются вместе, а в статистике они считаются связанными только в том случае, если A встречается относительно чаще среди B, чем среди не-B. Если 90% курящих страдают плохим пищеварением, то мы не можем сказать, что курение и плохое пищеварение связаны, пока не будет показано, что среди некурящих страдают плохим пищеварением менее чем 90%.

Стандартные ошибки коэффициентов

33.8 Теперь мы будем рассматривать таблицу 2×2 как выборку и найдем в предположении независимости стандартные ошибки коэффициентов, определенных в **33.6**.

Для обозначения дифференциала мы воспользуемся символом ∂ вместо d, поскольку последний уже использован нами. Получаем

$$\frac{\partial e}{e} = \frac{\partial b}{b} + \frac{\partial c}{c} - \frac{\partial a}{a} - \frac{\partial d}{d},$$

откуда

$$\frac{\mathrm{D}e}{e^2} = \sum_u \frac{\mathrm{D}u}{u^2} + 2 \sum_{u,\,v} \left\{ \pm \frac{\operatorname{cov}(u, v)}{uv} \right\}, \qquad (33.13)$$

*) Если следовать этому соглашению, то, вообще говоря, следует отличать полную связь между A и B от полной связи между B и A.

где u и v принимают значения a, b, c, d. Мы воспользуемся результатами, относящимися к мультиномиальному распределению, для которого

$$\mathrm{D}a = \frac{a(n-a)}{n}, \quad \operatorname{cov}(a, b) = -\frac{ab}{n}.$$

Подставляя эти значения в (33.13), находим

$$\mathrm{D}e = e^2\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right). \tag{33.14}$$

Отсюда легко получаем

$$\mathrm{D}Q = \tfrac{1}{4}(1-Q^2)^2\left\{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right\}, \tag{33.15}$$

$$\mathrm{D}Y = \tfrac{1}{16}(1-Y^2)^2\left\{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right\}. \tag{33.16}$$

Выборочная дисперсия V может быть найдена аналогичным образом, только выкладки будут длиннее. Мы получим

$$V = \frac{1}{n}\left[1 - V^2 + \left(V + \frac{1}{2}V^3\right)\frac{(a-d)^2-(b-c)^2}{\{(a+b)(a+c)(b+d)(c+d)\}^{1/2}} - \right.$$
$$\left. - \frac{3}{4}V^2\left\{\frac{(a+b-c-d)^2}{(a+b)(c+d)} - \frac{(a+c-b-d)^2}{(a+c)(b+d)}\right\}\right]. \tag{33.17}$$

В этих формулах предполагается, как это обычно делается в теории больших выборок, что в выборочных дисперсиях могут быть использованы сами наблюдаемые частоты вместо их математических ожиданий.

Частная связь

33.9 Введенные выше коэффициенты характеризуют связь двух признаков в статистическом смысле, но для того, чтобы выяснить, обусловлена ли эта связь некоторой причинной зависимостью, часто бывает необходимо, так же как это делалось для корреляции в **27.1**, рассмотреть связь в некоторых подсовокупностях. Предположим, например, что замечена положительная связь между прививкой и отсутствием заболевания. Напрашивается вывод, что прививка предохраняет от заболевания, но это не обязательно так. Возможно, что прививки чаще делаются людям, которые принадлежат к более богатым классам и живут в лучших гигиенических условиях и которые, тем самым, труднее поддаются заболеванию или меньше подвержены риску. Другими словами, связь между A и B может быть обусловлена связью обоих этих признаков с третьим признаком C (благосостоянием). Поэтому мы рассмотрим связь A и B при условии, что C фиксировано.

Связи в подсовокупностях называются *частными связями*. По аналогии с (33.6) A и B называются положительно связанными в совокупности C, если

$$(ABC) > \frac{(AC)(BC)}{(C)}, \tag{33.18}$$

где, например, (ABC) — число членов, обладающих признаками A, B и C. Мы можем также определить коэффициенты частной связи, коллигации и т. п.; например, коэффициент частной связи определяется равенством

$$Q_{AB.C} = \frac{(ABC)(\alpha\beta C) - (A\beta C)(\alpha BC)}{(ABC)(\alpha\beta C) + (A\beta C)(\alpha BC)}, \tag{33.19}$$

которое получается из (33.9) добавлением C ко всем символам, представляющим частоты.

Пример 33.2

Следующий пример, хотя и не характерный для современной генетической теории, представляет исторический интерес, так как показывает ранние попытки применения численных методов к изучению наследственности.

В *Natural Inheritance* Гальтона собраны сведения о цвете глаз родителей и детей для 78 семей, содержащих не менее шести братьев или сестер. Обозначив светлоглазого ребенка A, светлоглазого родителя B, а светлоглазого прародителя C, просмотрим каждое поколение и проследим, имеет ли светлоглазый ребенок светлоглазого родителя и светлоглазого прародителя. Число таких детей обозначим (ABC). Проведем такой подсчет также для других сочетаний признаков. Например, символ $(A\beta\gamma)$ будет обозначать число светлоглазых детей, имеющих несветлоглазых родителей и несветлоглазых прародителей. В результате для восьми возможных групп было получено:

$$\begin{aligned} (ABC) &= 1928, & (\alpha BC) &= 303, \\ (AB\gamma) &= 596, & (\alpha B\gamma) &= 225, \\ (A\beta C) &= 552, & (\alpha\beta C) &= 395, \\ (A\beta\gamma) &= 508, & (\alpha\beta\gamma) &= 501. \end{aligned}$$

Сначала обсудим вопрос: существует ли какая-либо связь между родителями и ребенком в отношении цвета глаз? Рассмотрим группу прародитель — родитель (связь между B и C) и группу родитель — ребенок (связь между A и B).

Доля светлоглазых среди детей светлоглазых прародителей равна $(BC)/(C) = 2231/3178 = 70{,}2\%$. Доля светлоглазых среди детей несветлоглазых прародителей равна $(B\gamma)/(\gamma) = 821/1830 = 44{,}9\%$. Аналогично $(AB)/(B) = 82{,}7\%$ и $(AB)/(\beta) = 54{,}2\%$. Очевидно, что в этой совокупности наблюдений существует по-

ложительная связь между родителями и ребенком в отношении цвета глаз.

Теперь рассмотрим соотношение между цветом глаз прародителей и внуков. Доля светлоглазых среди внуков светлоглазых прародителей равна $(AC)/(C) = 2480/3178 = 78{,}0\%$, а среди внуков несветлоглазых прародителей $(A\gamma)/(\gamma) = 1104/1830 = 60{,}3\%$.

Таким образом, связь между цветом глаз прародителей и внуков также положительна. Можно представить эти данные в виде таблиц:

Признак	A	α	Сумма
B	2524	528	3052
β	1060	896	1956
Сумма	3584	1424	5008

Признак	C	γ	Сумма
B	2231	821	3052
β	947	1009	1956
Сумма	3178	1830	5008

Признак	C	γ	Сумма
A	2480	1104	3584
α	698	726	1424
Сумма	3178	1830	5008

Коэффициенты связи и коллигации Q и Y равны:

	Q	Y
Прародители — родители .	0,487	0,260
Родители — дети	0,603	0,336
Прародители — внуки . . .	0,401	0,209

Теперь возникает вопрос: определяется ли сходство между прародителем и внуком только сходством между прародителем и родителем и между родителем и ребенком? Чтобы это исследовать, мы рассмотрим связи между прародителем и внуком в частных совокупностях «светлоглазые родители» и «несветлоглазые родители», т. е. связи между A и C для B и β.

Если рассматривать только светлоглазых родителей, то доля светлоглазых среди внуков светлоглазых прародителей равна $(ABC)/(BC) = 1938/2231 = 86{,}4\%$, тогда как доля светлоглазых среди внуков несветлоглазых прародителей $(AB\gamma)/(B\gamma) = 596/821 = 72{,}6\%$.

Если же рассматривать только несветлоглазых родителей, то доля светлоглазых среди внуков светлоглазых прародителей равна $(A\beta C)/(\beta C) = 552/947 = 58{,}3\%$, а доля светлоглазых среди внуков несветлоглазых прародителей равна $(A\beta\gamma)/(\beta\gamma) = 508/1009 = 50{,}3\%$.

В обоих случаях имеется положительная частная связь. Таким образом, связь между прародителями и внуками не может быть полностью обусловлена связью между прародителями и

родителями и связью между родителями и детьми. Полученные результаты были истолкованы как указывающие на существование наследственной связи между прародителями и внуками, помимо наследственной связи между родителями и детьми.

Приведем соответствующие таблицы:

Таблица 33.1

Светлоглазые родители					Несветлоглазые родители				
Прародители					Прародители				
Внуки		BC	$B\gamma$	Сумма	Внуки		βC	$\beta\gamma$	Сумма
	AB	1928	596	2524		$A\beta$	552	508	1060
	αB	303	225	428		$\alpha\beta$	395	501	896
	Сумма	2231	821	3052		Сумма	947	1009	1956

Коэффициенты связи и коллигации равны:

$$Q_{AC\cdot B} = 0{,}412, \quad Q_{AC\cdot\beta} = 0{,}159,$$

$$Y_{AC\cdot B} = 0{,}216, \quad Y_{AC\cdot\beta} = 0{,}080.$$

33.10 Если рассматривается p различных признаков, то число частных связей может быть очень большим даже для умеренных значений p. Мы можем, например, выбрать два признака $\binom{p}{2}$ способами и рассматривать их связи во всех возможных подсовокупностях, определяемых остальными $p-2$ признаками; число таких подсовокупностей равно 3^{p-2}. Таким образом, мы получаем $\binom{p}{2}\cdot 3^{p-2}$ связей.

В действительности одной из главных трудностей в обработке данных, характеризующихся множественной дихотомией (или, более того, множественной политомией), является наличие большого числа таблиц с результатами.

Полезно привести в связи с этим следующий результат. Обобщая определение D в (33.8), мы получаем

$$D_{AB\cdot C} + D_{AB\cdot\gamma} = \left\{(ABC) - \frac{(AC)(BC)}{(C)}\right\} + \left\{(AB\gamma) - \frac{(A\gamma)(B\gamma)}{(\gamma)}\right\} =$$

$$= (AB) - \frac{(A)(B)}{n} -$$

$$- \frac{n}{(C)(\gamma)}\left\{(AC)(BC) - \frac{(A)(C)(BC)}{n} - \frac{(B)(C)(AC)}{n} + \frac{(A)(B)(C)^2}{n^2}\right\} =$$

$$= D_{AB} - \frac{n}{(C)(\gamma)}\left\{(AC) - \frac{(A)(C)}{n}\right\}\left\{(BC) - \frac{(B)(C)}{n}\right\} =$$

$$= D_{AB} - \frac{n}{(C)(\gamma)} D_{AC} D_{BC}. \quad (33.20)$$

Таким образом, если A и B независимы в обеих совокупностях (C) и (γ), то $D_{AB\cdot C} = D_{AB\cdot\gamma} = 0$ и (33.20) дает

$$D_{AB} = \frac{n}{(C)(\gamma)} D_{AC} D_{BC}, \tag{33.21}$$

т. е. A и B не будут независимы во всей совокупности, за исключением случая, когда C не зависит от A или от B или от обоих признаков в этой совокупности. Сравните этот результат с (29.67) для частных корреляций, где из $\rho_{12.3} = 0$ вытекает, что $\rho_{12} = 0$, только если $\rho_{13} = 0$ или $\rho_{23} = 0$.

Полученный результат показывает, что при объединении двух совокупностей (C) и (γ) могут возникнуть кажущиеся связи, а действительные связи при этом могут оказаться скрытыми. Если A и C, B и C связаны, то из (33.20) получаем

$$D_{AB} = \frac{n}{(C)(\gamma)} D_{AC} D_{BC} + D_{AB\cdot C} + D_{AB\cdot\gamma}, \tag{33.22}$$

так что если A и B положительно связаны в (C) и отрицательно связаны в (γ), то D_{AB} может обратиться в нуль; это означает, что A и B предстанут независимыми во всей совокупности.

Пример 33.3

Рассмотрим случай, когда одни больные подвергались лечению, а другие — нет. Пусть A означает выздоровление, а B — лечение, и пусть таблица 2×2 имеет частоты:

	B	β	Сумма
A	100	200	300
α	50	100	150
Сумма	150	300	450

Здесь $(AB) = 100 = \frac{(A)(B)}{n}$, т. е. признаки независимы. Отсюда можно сделать вывод, что лечение нисколько не способствует выздоровлению.

Обозначим лиц мужского пола C, а лиц женского пола γ. Пусть частоты среди мужчин и женщин представлены таблицами:

Мужчины

	BC	βC	Сумма
AC	80	100	180
αC	40	80	120
Сумма	120	180	300

Женщины

	$B\gamma$	$\beta\gamma$	Сумма
$A\gamma$	20	100	120
$\alpha\gamma$	10	20	30
Сумма	30	120	150

В группе мужчин получаем

$$Q_{AB.C} = \frac{(80 \times 80) - (100 \times 40)}{(80 \times 80) + (100 \times 40)} = 0{,}231,$$

а в группе женщин

$$Q_{AB.\bar{C}} = -0{,}429.$$

Теперь видно, что лечение и выздоровление положительно связаны в группе мужчин и отрицательно связаны в группе женщин. Кажущаяся независимость в объединенной группе возникает из-за того, что обе эти связи взаимно уничтожаются.

Встречаются еще более парадоксальные случаи. Две таблицы, имеющие связи одного знака, могут при объединении дать таблицу со связью противоположного знака. Читатель может попробовать самостоятельно построить такой пример.

Вероятностная интерпретация мер связи

33.11 Теория рассмотренных нами в **33.5—10** мер связи, начавшая развиваться после 1900 г., старалась описать силу связи одним исчерпывающим коэффициентом. Однако, как мы уже видели в **26.10** на примере коэффициента корреляции, не всегда бывает разумно предполагать, что связь может быть адекватно описана одним коэффициентом. Гудмэн и Крускал (1954, 1959), подробно изложившие историю развития мер связи и составившие очень полную библиографию, привели веские аргументы в пользу выбора меры, допускающей интерпретацию в терминах решаемой задачи. Они рассматривают в основном политомические таблицы, т. е. таблицы с двумя или более классами по строкам и по столбцам. Мы вернемся к этому позже, в **33.35** и **33.40**, когда будем рассматривать политомию. Сейчас, однако, отметим, что при некоторых условиях коэффициентам Q и Y, определенным для таблицы 2×2, можно дать операционную интерпретацию.

33.12 Рассмотрим случайный выбор двух индивидов из совокупности n индивидов, классифицированных в соответствии с таблицей (33.1), так что каждый индивид попадает в одну из четырех категорий таблицы. Пусть для $i = 1, 2$

$$a_i = \begin{cases} +1, & \text{если индивид обладает признаком } A, \\ 0 & \text{в противном случае,} \end{cases}$$

$$b_i = \begin{cases} +1, & \text{если индивид обладает признаком } B, \\ 0 & \text{в противном случае,} \end{cases}$$

$$a_0 = a_1 - a_2, \qquad b_0 = b_1 - b_2.$$

Определим вероятности

$$\pi_s = \mathsf{P}\{a_0 b_0 = 1\}, \quad \pi_d = \mathsf{P}\{a_0 b_0 = -1\}, \quad \pi_t = \mathsf{P}\{a_0 b_0 = 0\}.$$

Тогда коэффициент

$$\gamma = \frac{\pi_s - \pi_d}{1 - \pi_t} \tag{33.23}$$

равен вероятности $\pi_s/(1-\pi_t)$ того, что из двух индивидов, выбранных из популяции, один принадлежит к категории AB, а другой — к категории $\alpha\beta$, минус вероятность $\pi_d/(1-\pi_t)$ того, что один из индивидов принадлежит к категории $A\beta$, а другой — к αB. Используя обозначения (33.2), легко показать, что

$$\gamma = (ad - bc)/(ad + bc) \equiv Q$$

(Q было определено в (33.9)). Таким образом, Q имеет прямой вероятностный смысл, указанный выше.

33.13 Аналогичным образом рассмотрим случайный выбор одного индивида из популяции и предположим, что требуется (не зная ничего априори) угадать, обладает индивид признаком A или не-A. Мы получим наилучшую оценку, если предположим, что индивид принадлежит к группе, имеющей бóльшую из частот (A) и (α) в (33.1); вероятность того, что наша оценка правильна, будет согласно (33.2) равна

$$\rho_{m\cdot} = \frac{1}{n}\max(a+b,\ c+d).$$

Аналогично, если надо угадать, обладает индивид признаком B или не-B, то мы выберем группу, имеющую бóльшую из частот (B) и (β); вероятность успеха при этом будет равна

$$\rho_{\cdot m} = \frac{1}{n}\max(a+c,\ b+d).$$

Таким образом, если нам надо в половине случаев угадать категорию A, а в другой половине случаев категорию B, то вероятность успеха будет равна $\frac{1}{2}(\rho_{m\cdot} + \rho_{\cdot m})$, а вероятность ошибки

$$\pi_0 = 1 - \frac{1}{2}(\rho_{m\cdot} + \rho_{\cdot m}). \tag{33.24}$$

Допустим теперь, что мы знаем категорию B индивида и нам надо угадать категорию A. Наилучшая оценка получится, если мы выберем наибольшую по численности категорию в соответствующем столбце таблицы. Вероятность успеха будет тогда равна $\frac{\max(a,\ c)}{a+c}$ или $\frac{\max(b,\ d)}{b+d}$ для соответствующих столбцов. Поскольку эти столбцы будут появляться в случайной выборке

с вероятностями $(a+c)/n$ и $(b+d)/n$ соответственно, то вероятность правильного оценивания категории A при заданной категории B будет равна

$$\pi_A = \frac{1}{n}\{\max(a, c) + \max(b, d)\}. \tag{33.25}$$

Аналогично, вероятность правильного оценивания категории B при заданной категории A равна

$$\pi_B = \frac{1}{n}\{\max(a, b) + \max(c, d)\}. \tag{33.26}$$

Если, как и в рассмотренном выше случае отсутствия информации, нам надо угадывать категории попеременно, то вероятность успеха будет равна среднему из (33.25) и (33.26), а вероятность ошибки будет равна

$$\pi_1 = 1 - \frac{1}{2n}\{\max(a, b) + \max(a, c) + \max(b, d) + \max(c, d)\}. \tag{33.27}$$

33.14 Теперь, используя (33.24) и (33.27), мы введем коэффициент

$$\lambda = \frac{\pi_0 - \pi_1}{\pi_0}, \tag{33.28}$$

характеризующий относительное уменьшение вероятности ошибки, вызванное знанием одной категории при предсказывании другой. Очевидно, $\pi_0 \geqslant \pi_1$, так что

$$0 \leqslant \lambda \leqslant 1. \tag{33.29}$$

Юл (1912) считал, что значение «разумной» меры связи не должно меняться, если каждую строку и каждый столбец таблицы 2×2 отдельно умножить на произвольные положительные константы. Используя этот принцип инвариантности, умножим строки и столбцы таблицы (33.2) на следующие константы:

первая строка: $(cd)^{1/2}/n$,
вторая строка: $(ab)^{1/2}/n$,
первый столбец: $(bd)^{1/2}/n$,
второй столбец: $(ac)^{1/2}/n$.

Отбрасывая общий для всех частот множитель, преобразуем таблицу (33.2) к виду

$$\begin{array}{cc|c} (ad)^{1/2} & (bc)^{1/2} & \frac{1}{2}m \\ (bc)^{1/2} & (ad)^{1/2} & \frac{1}{2}m \\ \hline \frac{1}{2}m & \frac{1}{2}m & m \end{array} \tag{33.30}$$

После некоторых очевидных преобразований можно показать, что любой коэффициент связи, удовлетворяющий принципу инвариантности, должен зависеть только от «перекрестного отношения» ad/bc. Меры Q и Y из (33.9), (33.10) удовлетворяют, очевидно, этому условию, а мера V из (33.12) не удовлетворяет. Эдвардс (1963) выводит это условие из других предположений.

Принцип инвариантности Юла позволяет нам связать λ из (33.28) с Y. Допустим пока, что выполняется неравенство $ad > bc$. В соответствии с (33.28) мы получаем для преобразованной таблицы (33.30)

$$\lambda = \frac{\frac{1}{2} - \left[1 - \frac{2}{m}(ad)^{1/2}\right]}{1/2} = \frac{2(ad)^{1/2} - \frac{1}{2}m}{\frac{1}{2}m}.$$

Но так как $(ad)^{1/2} + (bc)^{1/2} = \frac{1}{2}m$, то

$$\lambda = \frac{(ad)^{1/2} - (bc)^{1/2}}{(ad)^{1/2} + (bc)^{1/2}}. \tag{33.31}$$

Выражение (33.31) совпадает с определением Y в (33.10), однако приняв, что $ad > bc$, мы произвольно выбрали знак Y. Таким образом, вообще

$$\lambda \equiv |Y|, \tag{33.32}$$

что придает вероятностный смысл величине Y.

Асимптотические критерии независимости в таблицах 2×2

33.15 Мы будем считать теперь, что наблюденные частоты в таблице 2×2 получены в результате случайного выбора, и предположим, что в генеральной совокупности истинные вероятности, соответствующие частотам a, b, c, d, равны p_{11}, p_{12}, p_{21}, p_{22} соответственно. Будем записывать эти вероятности в виде таблицы:

p_{11}	p_{12}	$p_{1.}$
p_{21}	p_{22}	$p_{2.}$
$p_{.1}$	$p_{.2}$	1

(33.33)

где $p_{1.} = p_{11} + p_{12}$ и т. д. Предполагается, что выборка производится с возвращением (или, что то же самое, что генеральная совокупность бесконечна).

Таблицу (33.2) мы перепишем, используя симметричные обозначения:

n_{11}	n_{12}	$n_{1\cdot}$
n_{21}	n_{22}	$n_{2\cdot}$
$n_{\cdot 1}$	$n_{\cdot 2}$	n

Распределение выборочных частот дается мультиномиальным распределением с общим членом

$$L = \frac{n!}{n_{11}!\, n_{12}!\, n_{21}!\, n_{22}!}\, p_{11}^{n_{11}} p_{12}^{n_{12}} p_{21}^{n_{21}} p_{22}^{n_{22}}. \tag{33.34}$$

Чтобы оценить p_{ij}, найдем максимум правдоподобия по p_{ij} при ограничении $\sum p_{ij} = 1$. Если λ — множитель Лагранжа, то мы получаем

$$\frac{n_{11}}{p_{11}} - \lambda = 0, \qquad \text{или} \qquad n_{11} = \lambda p_{11},$$

и аналогично для трех других вероятностей. Суммируя все уравнения, находим $\lambda = n$. Оценки p_{ij} имеют вид

$$\hat{p}_{ij} = n_{ij}/n. \tag{33.35}$$

Это как раз то, чего следовало ожидать. Полученные оценки несмещены.

Известно, и мы этим уже пользовались в **33.8**, что дисперсии и ковариации n_{ij} имеют вид

$$\mathsf{D} n_{ij} = n p_{ij}(1 - p_{ij}), \qquad \operatorname{cov}(n_{ij}, n_{kl}) = -n p_{ij} p_{kl}.$$

Это точные результаты. Мы также знаем (см. пример 15.3), что совместное распределение n_{ij} стремится к многомерному нормальному распределению с такими дисперсиями и ковариациями. Теперь мы можем также отметить, что асимптотическая многомерная нормальность следует из того факта, что полученные оценки являются МП-оценками и что они удовлетворяют условиям пункта **18.26**.

33.16 Предположим теперь, что мы хотим проверить в таблице 2×2 гипотезу независимости

$$H_0\colon\ p_{11} p_{22} = p_{12} p_{21}. \tag{33.36}$$

Эта гипотеза, разумеется, является сложной, с одним ограничением и двумя степенями свободы. Будем считать p_{11} и p_{12} независимыми переменными и выразим через них p_{21} и p_{22}:

$$p_{21} = \frac{p_{11}(1 - p_{11} - p_{12})}{p_{11} + p_{12}}, \quad p_{22} = \frac{p_{12}(1 - p_{11} - p_{12})}{p_{11} + p_{12}}. \tag{33.37}$$

Если пренебречь константами, то логарифм функции правдоподобия, следовательно, равен

$$\log L = n_{11}\log p_{11} + n_{12}\log p_{12} + n_{21}\log p_{21} + n_{22}\log p_{22} =$$
$$= n_{11}\log p_{11} + n_{12}\log p_{12} +$$
$$+ n_{21}\{\log p_{11} + \log(1 - p_{11} - p_{12}) - \log(p_{11} + p_{12})\} +$$
$$+ n_{22}\{\log p_{12} + \log(1 - p_{11} - p_{12}) - \log(p_{11} + p_{12})\} =$$
$$= n_{.1}\log p_{11} + n_{.2}\log p_{12} + n_{2.}\{\log(1 - p_{1.}) - \log p_{1.}\}.$$

Чтобы оценить параметры, напишем уравнения

$$0 = \frac{\partial \log L}{\partial p_{11}} = \frac{n_{.1}}{p_{11}} - n_{2.}\left\{\frac{1}{1 - p_{1.}} + \frac{1}{p_{1.}}\right\} = \frac{n_{.1}}{p_{11}} - \frac{n_{2.}}{p_{1.}(1 - p_{1.})}, \quad (33.38)$$

$$0 = \frac{\partial \log L}{\partial p_{12}} = \frac{n_{.2}}{p_{12}} - \frac{n_{2.}}{p_{1.}(1 - p_{1.})}, \quad (33.39)$$

откуда находим МП-оценки, соответствующие гипотезе H_0,

$$\hat{\hat{p}}_{11} = \frac{n_{1.}}{n}\frac{n_{.1}}{n}, \qquad \hat{\hat{p}}_{12} = \frac{n_{1.}}{n}\frac{n_{.2}}{n}. \quad (33.40)$$

Соотношения (33.37) приводят к аналогичным выражениям для $\hat{\hat{p}}_{21}$ и $\hat{\hat{p}}_{22}$. Таким образом, вероятности оцениваются через произведения соответствующих маргинальных относительных частот. Это оправдывает определения связи путем сравнения с такими произведениями в **33.4—5**.

Подставляя найденные МП-оценки в ФП, получаем

$$L(n_{ij} \mid H_0, \hat{\hat{p}}_{ij}) \propto (n_{1.}n_{.1})^{n_{11}}(n_{1.}n_{.2})^{n_{12}}(n_{2.}n_{.1})^{n_{21}}(n_{2.}n_{.2})^{n_{22}}/n^{2n}, \quad (33.41)$$

тогда как безусловный максимум ФП достигается при подстановке оценок (33.35), что дает

$$L(n_{ij} \mid \hat{p}_{ij}) \propto n_{11}^{n_{11}}n_{12}^{n_{12}}n_{21}^{n_{21}}n_{22}^{n_{22}}/n^n. \quad (33.42)$$

Соотношения (33.41), (33.42) приводят к статистике критерия отношения правдоподобия

$$l = \left(\frac{n_{1.}n_{.1}}{nn_{11}}\right)^{n_{11}}\left(\frac{n_{1.}n_{.2}}{nn_{12}}\right)^{n_{12}}\left(\frac{n_{2.}n_{.1}}{nn_{21}}\right)^{n_{21}}\left(\frac{n_{2.}n_{.2}}{nn_{22}}\right)^{n_{22}}. \quad (33.43)$$

Вводя обозначения $\hat{\hat{np}}_{ij} = n_{i.}n_{.j}/n = e_{ij}$, можно переписать (33.43) в виде

$$l = \left(\frac{e_{11}}{n_{11}}\right)^{n_{11}}\left(\frac{e_{12}}{n_{12}}\right)^{n_{12}}\left(\frac{e_{21}}{n_{21}}\right)^{n_{21}}\left(\frac{e_{22}}{n_{22}}\right)^{n_{22}}. \quad (33.44)$$

33.17 Из общего результата пункта **24.7** следует, что $-2\log l$ имеет асимптотически распределение χ^2 с одной степенью свободы. Это легко показать следующим образом. Обозначая $D_{ij} = n_{ij} - e_{ij}$ (см. (33.8)) и разлагая $-2\log l$ до членов порядка $D^2 (= D_{ij}^2$ для всех $i, j)$, получим

$$-2\log l = 2\sum_{i=1}^{2}\sum_{j=1}^{2} e_{ij}\left(1+\frac{D_{ij}}{e_{ij}}\right)\left(\frac{D_{ij}}{e_{ij}} - \frac{1}{2}\frac{D^2}{e_{ij}^2}\right) = D^2\sum_i\sum_j \frac{1}{e_{ij}}. \tag{33.45}$$

(33.45) можно переписать в виде

$$-2\log l = \sum_i\sum_j \frac{(n_{ij}-e_{ij})^2}{e_{ij}} \equiv X^2. \tag{33.46}$$

Таким образом, мы для одного частного случая доказали асимптотическую эквивалентность критерия отношения правдоподобия и критерия согласия X^2, о которой мы говорили в **30.5**. Равенство (33.46) можно было бы получить сразу, замечая, что сложной гипотезе H_0 соответствует набор гипотетических частот e_{ij} и что проверка независимости равнозначна проверке согласия наблюдений с этими гипотетическими частотами. Как и в **30.10**, число степеней свободы равно числу классов (4) без единицы минус число оцениваемых параметров (2), т. е. единице.

Легко показать, что статистика X^2, определенная в (33.46), тождественно равна nV^2, где V — мера связи, определенная в (33.12). Мы предоставляем читателю проверить это.

Точный критерий независимости: модели для таблицы 2×2

33.18 Критерии независимости, полученные в **33.15—17**, являются асимптотическими относительно объема выборки n. Прежде чем строить точные критерии независимости для таблицы 2×2, мы должны обратить внимание на некоторые особенности, впервые замеченные Барнардом (1947a, b) и Пирсоном (1947).

Напомним, что математические ожидания в клетках таблицы 2×2 при гипотезе независимости двух категоризованных переменных равны

$$e_{ij} = \frac{n_{i\cdot}n_{\cdot j}}{n}, \quad i, j = 1, 2, \tag{33.47}$$

т. е. зависят только от четырех маргинальных частот и от объема выборки n. Поскольку сейчас мы хотим получить точные результаты, мы должны ясно представлять, каким путем была образована таблица и как, в частности, были получены маргиналь-

ные частоты. Даже при фиксированном n мы имеем три возможности в отношении маргинальных частот. Оба набора маргинальных частот могут быть случайными величинами, как это имеет место в случае, когда выборка объема n берется из двумерного распределения и затем делится, образуя двойную дихотомию. В другом случае один набор маргинальных частот может быть фиксирован, поскольку такая классификация состоит просто в различении двух выборок (скажем, мужчины и женщины), которые нужно сравнивать по другим признакам (скажем, число заболевших и не заболевших какой-то болезнью). Если объемы обеих выборок определены заранее (например, если мы хотим исследовать заданное число мужчин и женщин), то мы получим один фиксированный набор маргинальных частот и один переменный. Такое сравнение двух (или более) выборок по некоторому признаку часто называется проверкой *однородности* двух (или k) выборок.

Наконец, возможен третий случай, когда оба набора маргинальных частот фиксированы заранее. На практике он встречается гораздо реже, чем два других. Читатель может прежде, чем продолжить чтение, попытаться самостоятельно придумать такую ситуацию. Классический пример (см. Фишер (1935а)) относится к психофизическому эксперименту: субъект (человек) подвергается n испытаниям на различение двух объектов (например, масла и маргарина по вкусу). Каждый объект предлагается некоторое число раз (не обязательно одинаковое для обоих объектов), и субъекту сообщаются эти числа. Разумный субъект будет в этом случае стараться, чтобы маргинальные частоты его утверждений («масло» или «маргарин») совпали с действительными частотами.

Пример 33.4

Чтобы лучше понять материал пункта **33.18**, рассмотрим несколько примеров. Таблица приведенного выше примера 33.1, конечно, не относится к последнему типу с двумя фиксированными наборами маргинальных частот, однако без дополнительной информации непонятно, к какому из двух других типов она принадлежит. Возможно, были исследованы 818 человек и распределены в соответствии с таблицей 2×2. С другой стороны, возможно, что были проверены отдельно 279 сделавших прививку и 539 не сделавших прививку и классифицированы на «заболевших» и «незаболевших». Можно также разделить две выборки, включающие 69 заболевших и 749 незаболевших, на «сделавших прививку» и «не сделавших прививку». Таким образом, существует три пути построения такой таблицы: один — двойная дихотомия и два других — соответствующие проверки однородно-

сти. Для того чтобы сделать выбор между этими тремя вариантами, необходимо знать, как в действительности проводились наблюдения.

Для иллюстрации последнего типа, описанного в **33.18**, мы приводим искусственно составленную таблицу, относящуюся к эксперименту по различению масла и маргарина.

		Идентификация, сделанная субъектом: Масло	Маргарин	
Объект, представленный в действительности	Масло	4	11	15
	Маргарин	11	14	25
		15	25	40

33.19 У нас нет никаких оснований ожидать, что один и тот же метод анализа может подойти для трех различных ситуаций, описанных в **33.18** (хотя ниже, в **33.24**, мы увидим, что если речь идет о критерии независимости, то критерий для случая I будет оптимальным в двух других случаях). Дадим теперь вероятностную формулировку этих трех ситуаций. Начнем с самого простого случая, когда обе маргинальные частоты зафиксированы.

Случай I. Обе маргинальные частоты фиксированы. В случае выполнения гипотезы

$$H_0\colon \frac{p_{11}}{p_{1\cdot}} = \frac{p_{21}}{p_{2\cdot}} \tag{33.48}$$

вероятность осуществления таблицы

n_{11}	n_{12}	$n_{1\cdot}$
n_{21}	n_{22}	$n_{2\cdot}$
$n_{\cdot 1}$	$n_{\cdot 2}$	n

(33.49)

при фиксированных маргинальных частотах равна

$$P_{\mathrm{I}} = \mathsf{P}\{n_{ij} \mid n,\ n_{1\cdot},\ n_{\cdot 1}\} = \mathsf{P}\{n_{ij} \mid n,\ n_{1\cdot}\}/\mathsf{P}\{n_{\cdot 1} \mid n\} =$$

$$= \binom{n_{1\cdot}}{n_{11}}\binom{n_{2\cdot}}{n_{21}}\Big/\binom{n}{n_{\cdot 1}} = \frac{n_{1\cdot}!\, n_{\cdot 1}!\, n_{2\cdot}!\, n_{\cdot 2}!}{n!\, n_{11}!\, n_{12}!\, n_{21}!\, n_{22}!}. \tag{33.50}$$

Формула (33.50) симметрична относительно частот и относительно маргинальных частот, как это и должно быть вследствие симметричности ситуации. Поскольку маргинальные частоты фиксированы, то только одна из частот n_{ij} может изменяться

независимо. Не ограничивая общности, можно предположить, что это n_{11}. Тогда (33.50) можно рассматривать как распределение n_{11}, которое оказывается гипергеометрическим (см. **5.18**). Действительно, (33.50) совпадает с гипергеометрической вероятностью (5.48), если произвести подстановки

$$N=n,\quad n=n_{1\cdot},\quad Np=n_{\cdot 1},\quad Nq=n_{\cdot 2},\quad N-n=n_{2\cdot},$$
$$j=n_{11},\quad n-j=n_{12},\quad Np-j=n_{21},\quad Nq-(n-j)=n_{22}.$$

Среднее и дисперсия n_{11} согласно (5.53) и (5.55) равны

$$\mathsf{M}(n_{11})=\frac{n_{1\cdot}n_{\cdot 1}}{n},\qquad \mathsf{D}(n_{11})=\frac{n_{1\cdot}n_{\cdot 1}n_{2\cdot}n_{\cdot 2}}{n^2(n-1)}. \tag{33.51}$$

Величина n_{11} распределена асимптотически нормально со средним и дисперсией (33.51). Таким образом, случайная величина

$$t=\frac{n_{11}-n_{1\cdot}n_{\cdot 1}/n}{\left\{\dfrac{n_{1\cdot}n_{\cdot 1}n_{2\cdot}n_{\cdot 2}}{n^2(n-1)}\right\}^{1/2}} \tag{33.52}$$

асимптотически имеет стандартное нормальное распределение. Заменяя $(n-1)$ на n, мы видим, что величина t эквивалентна $n^{1/2}V$ (V определено в (33.12)), следовательно, величина t^2 эквивалентна статистике X^2, определенной в (33.46). Это свидетельствует о том, что общий асимптотический критерий пункта **33.17** применим в данной ситуации.

33.20 С помощью (33.50) мы можем вычислить точную вероятность любого заданного набора частот. Суммируя вероятности, соответствующие «хвосту» распределения n_{11}, можно построить критическую область для точного критерия, впервые предложенного Фишером. Эта процедура проиллюстрирована на следующем примере.

Пример 33.5 (Данные из работы Иэйтса (1934), принадлежащие М. Хеллмэну).

В следующей таблице классифицированы 42 ребенка в зависимости от состояния их зубов и способа кормления.

	Нормальные зубы	Зубы с неправильным прикусом	Сумма
Кормление грудью	4	16	20
Искусственное кормление	1	21	22
Сумма	5	37	42

Очевидно, здесь не могут быть зафиксированы обе маргинальные частоты, однако в данный момент мы используем эти данные в иллюстративных целях, а позднее, в **33.24**, увидим, что это вполне оправдано.

Мы выберем в качестве n_{11} частоту с наименьшим диапазоном изменения, т. е. одну из двух частот, которым соответствуют наименьшие маргинальные частоты. В нашем случае такой (фиксированной) маргинальной частотой будет $n_{\cdot 1} = 5$, которой соответствует изменение n_{11} от 0 до 5.

Вероятность того, что $n_{11} = 0$, согласно (33.50) равна

$$\frac{5!\,37!\,20!\,22!}{42!\,20!\,0!\,5!\,17!} = 0{,}03096.$$

Вероятность того, что $n_{11} = 1, 2, \ldots$, легче всего получить, умножая это значение соответственно на

$$\frac{5 \times 20}{1 \times 18}, \quad \frac{4 \times 19}{2 \times 19}, \quad \frac{3 \times 18}{3 \times 20} \quad \text{и т. д.}$$

Эти вероятности равны:

Число детей, вскормленных грудью и имеющих нормальные зубы (n_{11})	Вероятность	Накопленная вероятность
0	0,0310	1,0001
1	0,1710	0,9691
2	0,3440	0,7971
3	0,3096	0,4531
4	0,1253	0,1435
5	0,0182	0,0182
	1,0001	

Для проверки гипотезы независимости против альтернативы, что нормальность зубов положительно связана с грудным кормлением, мы будем использовать критическую область, состоящую из больших значений n_{11} (число детей с нормальными зубами, вскормленных грудью). У нас есть выбор между двумя «разумными» значениями размера точного критерия. Для $\alpha = 0{,}0182$ только $n_{11} = 5$ приводит к отклонению гипотезы; для $\alpha = 0{,}1435$ к отклонению гипотезы приводит $n_{11} = 4$ или 5. Вероятно, большинство статистиков использовало бы первую критическую область, что в нашем случае ($n_{11} = 4$) соответствовало бы принятию гипотезы о независимости.

33.21 Для точного критерия, основанного на (33.50), имеются таблицы. Финни (1948) приводит значения n_{21} (у него b), необходимые для отклонения гипотезы о независимости для значений $n_{1\cdot}$, $n_{2\cdot}$ (или $n_{\cdot 1}$, $n_{\cdot 2}$) и n_{11}, не

превосходящих 15, и для односторонних критериев размера $\alpha \leqslant 0{,}05$, 0,025, 0,01, 0,005 с указанием точного размера в каждом случае. Таблица Финни перепечатана в *Biometrika Tables*. Лача (1953) продолжил таблицу Финни до $n_{1\cdot}$, $n_{2\cdot} = 20$ *). Эти таблицы продолжены до $n_{1\cdot}$, $n_{2\cdot} = 40$ Финни и др. (1963). Армсен (1955) построил таблицы для односторонних и двусторонних критериев размера $\alpha \leqslant 0{,}05$, 0,01 и $n \leqslant 50$. Бросс и Кастен (1957) построили графики для односторонних критериев размера $\alpha = 0{,}05$, 0,025, 0,01, 0,005 (или двусторонних критериев размера 2α) с минимальной маргинальной частотой (скажем, $n_{\cdot 1}$) $\leqslant 50$, основанные на аппроксимации гипергеометрического распределения (33.50) биномиальным $\binom{n_{1\cdot}}{n_{11}}\binom{n_{\cdot 1}}{n}^{n_{11}}\binom{n_{\cdot 2}}{n}^{n_{21}}$. Критические значения, даваемые графиками, являются консервативными, если $n_{\cdot 1}/n$ не очень мало.

Случай II. Одна маргинальная частота фиксирована; однородность.

33.22 Мы будем записывать гипотезу (являющуюся теперь гипотезой о равенстве вероятностей в двух совокупностях), как и прежде, в форме (33.48), но теперь $n_{1\cdot}$ и $n_{2\cdot}$ зафиксированы, а n_{11}, n_{21} — независимые случайные величины, так что частота $n_{\cdot 1}$ (и, следовательно, ее дополнение $n_{\cdot 2}$) — случайная величина. Мы проверяем гипотезу (33.48), рассматривая соответствующую разность отношений

$$u = \frac{n_{11}}{n_{1\cdot}} - \frac{n_{21}}{n_{2\cdot}}. \tag{33.53}$$

При гипотезе эта разность асимптотически нормальна со средним 0 и дисперсией

$$\mathsf{D}u = p(1-p)\left(\frac{1}{n_{1\cdot}} + \frac{1}{n_{2\cdot}}\right),$$

где p — гипотетическое общее значение $\frac{p_{11}}{p_{1\cdot}}$ и $\frac{p_{21}}{p_{2\cdot}}$.

Несмещенной оценкой для $p(1-p)$ служит объединенная оценка

$$\frac{n}{n-1}\left(\frac{n_{\cdot 1}}{n}\right)\left(\frac{n_{\cdot 2}}{n}\right),$$

так что оценка дисперсии u равна

$$\hat{\mathsf{D}}u = \frac{n_{\cdot 1}n_{\cdot 2}}{n(n-1)}\left(\frac{1}{n_{1\cdot}} + \frac{1}{n_{2\cdot}}\right) = \frac{n_{\cdot 1}n_{\cdot 2}}{(n-1)\,n_{1\cdot}n_{2\cdot}}.$$

Таким образом, случайная величина

$$t = \frac{u - \mathsf{M}(u)}{(\hat{\mathsf{D}}u)^{1/2}} = \frac{\dfrac{n_{11}}{n_{1\cdot}} - \dfrac{n_{21}}{n_{2\cdot}}}{\left\{\dfrac{n_{\cdot 1}n_{\cdot 2}}{(n-1)\,n_{1\cdot}n_{2\cdot}}\right\}^{1/2}} \tag{33.54}$$

*) Несколько расширенные по сравнению с указанными таблицы имеются в книге *Таблицы математической статистики*. (*Прим. ред.*)

распределена асимптотически нормально с нулевым средним и единичной дисперсией. Мы видим, что (33.54) совпадает с (33.52). Следовательно, асимптотические критерии совпадают в этих двух случаях.

Однако при малых выборках критерии различны. Когда справедлива гипотеза, мы имеем теперь

$$P_{\mathrm{II}} = \mathsf{P}\{n_{11} \mid p,\ n_{1\cdot}\}\mathsf{P}\{n_{21} \mid p,\ n_{2\cdot}\} =$$

$$= \binom{n_{1\cdot}}{n_{11}} p^{n_{11}}(1-p)^{n_{12}} \binom{n_{2\cdot}}{n_{21}} p^{n_{21}}(1-p)^{n_{22}}, \quad (33.55)$$

а это есть определенная в (33.50) вероятность P_{I}, умноженная на биномиальный коэффициент

$$\binom{n}{n_{\cdot 1}} p^{n_{\cdot 1}}(1-p)^{n_{\cdot 2}}.$$

Очевидно, что это должно быть так, потому что вероятности, соответствующие фиксированной частоте $n_{\cdot 1}$, умножаются теперь на вероятность самой $n_{\cdot 1}$. В отличие от (33.50), вероятность (33.55) зависит от неизвестного параметра p и поэтому не может быть вычислена.

Случай III. Маргинальные частоты не фиксированы; двойная дихотомия.

33.23 Теперь рассмотрим случай, когда n фиксировано, но ни одна из маргинальных частот не фиксирована. Гипотеза теперь состоит в двумерной независимости. Асимптотический критерий уже получен нами в **33.15—17**. Точная вероятность появления частоты n_{11} теперь равна

$$P_{\mathrm{III}} = \mathsf{P}\{n_{11} \mid n_{1\cdot},\ n_{\cdot 1},\ n\}\mathsf{P}\{n_{\cdot 1} \mid p,\ n\}\mathsf{P}\{n_{1\cdot} \mid p',\ n\}, \quad (33.56)$$

где p' — гипотетическое общее значение $p_{11}/p_{\cdot 1}$ и $p_{12}/p_{\cdot 2}$. Произведение первых двух сомножителей, стоящих в правой части (33.56), эквивалентно (33.55), а третий сомножитель равен

$$\mathsf{P}\{n_{1\cdot} \mid p',\ n\} = \binom{n}{n_{1\cdot}} (p')^{n_{1\cdot}}(1-p')^{n_{2\cdot}}. \quad (33.57)$$

Выражение (33.56) зависит, таким образом, от неизвестных параметров $(p,\ p')$, и его нельзя вычислить.

Оптимальный критерий для таблиц 2×2

33.24 Мы можем теперь доказать один замечательный результат (впервые полученный Точером (1950)), утверждающий, что точный критерий, основанный на вероятностях (33.50) случая I, в действительности является РНМ несмещенным критерием для случаев II и III. Доказывается он очень просто. Рас-

пределение n_{11} (33.55) для случая II, соответствующее гипотезе H_0, содержит единственный мешающий параметр — гипотетическое общее значение $p_{11}/p_{1\cdot}$ и $p_{21}/p_{2\cdot}$. Легко проверить, что когда гипотеза H_0 выполнена, объединенная оценка $\hat{p} = \dfrac{n_{11}+n_{21}}{n_{1\cdot}+n_{2\cdot}} = \dfrac{n_{\cdot 1}}{n}$ достаточна для p. В силу **23.10** она полна и ее распределение имеет линеаризованную экспоненциальную форму (23.17). Таким образом, согласно **23.30—31** односторонние и двусторонние РНМН критерии для H_0 должны быть основаны на условном распределении n_{11} при фиксированном $n_{\cdot 1}$, т. е. (поскольку $n_{1\cdot}$ уже фиксировано) на (33.50). См. упражнение 23.22.

Аналогично, в случае III имеется два мешающих параметра (p, p') в (33.56), для которых пара статистик $\left(\dfrac{n_{\cdot 1}}{n}, \dfrac{n_{1\cdot}}{n}\right)$ совместно достаточна и полна при гипотезе H_0. Таким образом, снова в силу **23.10** и **23.30—31** РНМН критерий для H_0 должен быть основан на условном распределении n_{11} при заданных $(n_{\cdot 1}, n_{1\cdot})$, т. е. на (33.50).

Итак, условное распределение случая I (все маргинальные частоты фиксированы) дает РНМН критерии как для случая однородности, так и для случая двойной дихотомии.

Следует заметить, что эти результаты выполняются строго, только если допускается рандомизация для получения критериев любого размера α; дискретность распределений в таблице 2×2 ограничивает выбор размеров критерия (см. пример 33.5, а также **20.22**). Если, однако, частоты не очень малы, то обычно для условного точного критерия, основанного на (33.50), имеется по крайней мере одно «разумное» значение α, так что трудность является скорее теоретической, нежели практической.

33.25 Хотя один и тот же критерий подходит для всех трех ситуаций, он может иметь разные функции мощности, поскольку альтернативы для независимости, очевидно, различны в этих трех ситуациях. Для случая II (однородность) Беннетт и Сюй (1960), пользуясь таблицами Финни и Лача (см. **33.21**), построили графики функции мощности. Патнайк (1948) получил приближения, пригодные для больших объемов выборок; см. также Хеннан и Харкнесс (1963). Пирсон и Меррингтон (1948) осуществили выборочные эксперименты с целью получения функции мощности точного и асимптотического критериев в случае I (все маргинальные частоты фиксированы). Харкнесс и Катц (1964) провели сравнение функций мощности для трех случаев; см. также Харкнесс (1965).

33.26 Берджер (1961) построил асимптотические χ^2-критерии для проверки равенства мер связи в двух таблицах 2×2. Им рассмотрены три меры связи: (а) отношение p_{11}/p_{12} (см. (33.33)); (б) коэффициент Q, определенный в (33.9); (в) мера

связи, эквивалентная коэффициенту V, определенному в (33.12). Гудмэн (1936а) дает обобщение на случай k таблиц 2×2, а также (1964а) приводит другие методы, основанные на перекрестном отношении ad/bc. См. также **33.62** ниже.

Поправка на непрерывность в асимптотическом χ^2-критерии

33.27 Для улучшения непрерывной аппроксимации (33.52) точного дискретного распределения (33.50) в случае I Иэйтс (1934) предложил поправку на непрерывность, аналогичную упомянутой в **31.80**. Эта поправка (при использовании (33.12)) состоит в том, что в выражении

$$X^2 \equiv nV^2 = \frac{n(ad-bc)^2}{(a+b)(b+c)(b+d)(c+d)} \tag{33.58}$$

член $(ad - bc)$ в числителе заменяется на $|ad - bc| - \frac{1}{2}n$, что сводится (если $ad > bc$) к увеличению b и c на 1/2 и уменьшению a и d на 1/2. Таким образом, статистика критерия с поправкой имеет вид

$$X_c^2 = \frac{n\left\{|ad-bc| - \frac{1}{2}n\right\}^2}{(a+b)(a+c)(b+d)(c+d)}. \tag{33.59}$$

Эффект от поправки рассматривается в примере 33.6.

Пример 33.6

В примере 33.5 мы нашли, что вероятность неравенства $n_{11} \geqslant 4$ равна 0,1435. Сравним этот результат с результатом, полученным с помощью асимптотического χ^2-распределения с одной степенью свободы. Пользуясь таблицей примера 33.5, получаем следующее значение для (33.58):

$$X^2 = 42\,\frac{\{4 \cdot 21 - 16 \cdot 1\}^2}{20 \cdot 22 \cdot 5 \cdot 37} = 2{,}386.$$

Из таблицы 4б Приложения находим $\mathsf{P}\{X^2 \geqslant 2{,}386\} = 0{,}122$ (более точное значение равно 0,1224). Мы нашли, однако, вероятность, связанную с двусторонним критерием, поскольку X^2 есть квадрат нормального отклонения. Для сравнения с точным критерием надо разделить 0,1224 на 2, что дает значение 0,0612, являющееся, конечно, довольно плохим приближением для точного значения 0,1435.

Если же мы воспользуемся поправкой на непрерывность, то согласно формуле (33.59) получим следующее подправленное значение X_c^2:

$$\frac{42 \cdot (68 - 21)^2}{20 \cdot 22 \cdot 5 \cdot 37} = 1{,}140.$$

Соответствующая вероятность согласно таблице χ^2-распределения равна 0,2854. Разделив 0,2854 на 2, мы получим значение 0,1427, являющееся отличным приближением для точного значения 0,1435.

Кокрэн (1954) считает, что величина X^2 с поправкой на непрерывность служит адекватным приближением для точного критерия при $n \geqslant 40$, а если ни одна из гипотетических частот не меньше 5, то при $n \geqslant 20$.

В случаях II и III поправка на непрерывность не улучшает приближения (33.52) для (33.55) и (33.56) (см. Плэкетт (1964)).

Лэнкастер (1949а) изучил эффект от поправок на непрерывность в случаях, когда складывается несколько значений X^2, соответствующих разным таблицам. В такой ситуации не следует применять поправку на непрерывность к каждому значению X^2, поскольку это может привести к значительному смещению. Лэнкастер показывает, что если исходные таблицы не могут быть объединены, то лучше всего сложить значения X^2, вычисленные без поправок на непрерывность. Аналогичные результаты для односторонних критериев получены Иэйтсом (1955).

Таблицы $r \times c$: измерение связи

33.28 Мы рассмотрим теперь более общую ситуацию, когда две переменные классифицированы на две или более категорий. Обобщая введенные обозначения, мы будем записывать таблицу $r \times c$ в виде

$$\begin{array}{cccc|c} n_{11} & n_{12} & \cdots & n_{1c} & n_{1\cdot} \\ n_{21} & n_{22} & \cdots & n_{2c} & n_{2\cdot} \\ \cdot & \cdot & & \cdot & \\ \cdot & \cdot & & \cdot & \\ \cdot & \cdot & & \cdot & \\ n_{r1} & n_{r2} & \cdots & n_{rc} & n_{r\cdot} \\ \hline n_{\cdot 1} & n_{\cdot 2} & \cdots & n_{\cdot c} & n \end{array} \qquad (33.60)$$

В более ранней литературе таблицы (33.60) называются *таблицами сопряженности*. Общие рассмотрения пунктов **33.1—2** относятся также и к этой более общей категоризации двух переменных.

Задача измерения связи в такой таблице представляет серьезные трудности, связанные с существом задачи. В главе 26 мы пришли к выводу, что в случае числовых вариант может оказаться невозможным описать сложную схему взаимозависимости одним коэффициентом. Это справедливо и в данной ситуации. Наиболее успешные шаги в этом направлении были связаны с более или менее явными предположениями относительно природы распределения исходных вариант.

33.29 Если бы две переменные в (33.60) были независимы, то частота в клетке на пересечении i-й строки и j-го столбца

была бы равна $n_{i.}n_{.j}/n$. Следовательно, отклонение от независимости в этой клетке таблицы измеряется величиной

$$D_{ij} = n_{ij} - n_{i.}n_{.j}/n, \tag{33.61}$$

являющейся обобщением (33.8). Мы можем определить коэффициент связи в терминах так называемой *квадратичной сопряженности* $\sum_{i,j} D_{ij}^2/(n_{i.}n_{.j})$ и записать обобщение (33.46) в виде

$$X^2 = \sum_{i,j} \frac{D_{ij}^2}{n_{i.}n_{.j}/n} \equiv n\left\{\sum_{i,j} \frac{n_{ij}^2}{n_{i.}n_{.j}} - 1\right\}. \tag{33.62}$$

Если выполнена гипотеза о независимости, то величина X^2 имеет асимптотически χ^2-распределение, что легко понять с точки зрения проверки согласия, как в **33.17**. Число степеней свободы в данном случае равно числу классов минус 1, минус число параметров, т. е.

$$(rc - 1) - (r - 1) - (c - 1) = (r - 1)(c - 1).$$

33.30 Сама величина X^2 не очень подходит в качестве меры связи, поскольку верхняя граница X^2 стремится к бесконечности при возрастании n. Следуя Карлу Пирсону (1904), мы определим коэффициент

$$P = \left(\frac{X^2}{n + X^2}\right)^{1/2}, \tag{33.63}$$

который будем называть *коэффициентом сопряженности* Пирсона. Основанием для его введения служит тот факт, что если двумерное нормальное распределение с параметром корреляции ρ представлено в виде таблицы сопряженности, то $P^2 \to \rho^2$ при увеличении числа категорий в таблице. Однако при конечных r и c коэффициент P имеет недостатки. Он обращается в нуль, как это и должно быть, при полной независимости, и наоборот, если $P = 0$, то $X^2 = 0$, так что каждое отклонение D_{ij} равно нулю. Очевидно, что $0 \leqslant P \leqslant 1$. Однако P не имеет одного и того же для всех случаев верхнего предела, т. е. не удовлетворяет одному из требований, сформулированных в **33.5**. Рассмотрим, например, «квадратную» таблицу с $r = c$, в которой отличны от нуля только элементы главной диагонали. Тогда $n_{i.} = n_{.i} = n_{ii}$ для всех i, и поэтому согласно (33.62)

$$X^2 = n(r - 1),$$

так что в соответствии с (33.63)

$$P = \left(\frac{r - 1}{r}\right)^{1/2}.$$

Таким образом, даже в случае *полной связи* (см. 33.7) значение P (достигающее в этом случае максимума) зависит от числа строк и столбцов в таблице.

Чтобы исправить указанный недостаток, Чупров предложил другую функцию от X^2:

$$T = \left\{ \frac{X^2}{n\,[(r-c)\,(c-1)]^{1/2}} \right\}^{1/2}, \tag{33.64}$$

которая, когда $r = c$, достигает значения $+1$ в случае полной связи, как это было выше, однако не обладает этим свойством, когда $r \neq c$. Действительно, легко показать точно так же, как и выше, что максимальное достижимое значение X^2 равно $n \min(r-1, c-1)$ (оно достигается, когда все частоты лежат на самой длинной диагонали таблицы), и, таким образом, достижимая верхняя граница для P равна

$$\left\{ \frac{\min(r-1, c-1)}{1+\min(r-1, c-1)} \right\}^{1/2},$$

тогда как для T она равна

$$\left\{ \frac{\min(r-1, c-1)}{[(r-1)\,(c-1)]^{1/2}} \right\}^{1/2} = \left\{ \frac{\min(r-1, c-1)}{\max(r-1, c-1)} \right\}^{1/2}.$$

Следуя Крамеру (1946), определим коэффициент

$$C = \left\{ \frac{X^2}{n \min(r-1, c-1)} \right\}^{1/2} = T \left\{ \frac{\max(r-1, c-1)}{\min(r-1, c-1)} \right\}^{1/2}, \tag{33.65}$$

который может всегда достигать $+1$. Очевидно, $C = T$ для квадратной таблицы и $C > T$ в остальных случаях, хотя разница не очень велика, если r и c отличаются не очень сильно. Заметим также, что

$$\frac{P^2}{T^2} = \frac{[(r-1)\,(c-1)]^{1/2}}{1+(X^2/n)},$$

так что при увеличении n следует ожидать, что P будет больше T, если выполнена гипотеза независимости (при которой математическое ожидание X^2 равно $(r-1)\,(c-1)$). Рзаность P—T часто бывает значительной. См. упражнение 33.4.

Пример 33.7

В следующей таблице (из Гилби (W. H. Gilby), *Biometrika* 8, 94) дано распределение 1725 школьников, классифицированных (1) в соответствии с качеством их одежды и (2) в соответствии с их умственными способностями. Во втором случае были использованы следующие градации: A — умственно отсталый, B — медлительный и тупой, C — тупой, D — медлительный, но умный, E — достаточно умный, F — явно способный, G — очень способный.

Таблица 33.2

Как одевается \ Способности	A и B	C	D	E	F	G	Сумма
Очень хорошо	33	48	113	209	194	39	636
Хорошо	41	100	202	255	138	15	751
Сносно	39	58	70	61	33	4	265
Очень плохо	17	13	22	10	10	1	73
Сумма	130	219	407	535	375	59	1725

Нас интересует связь между одеждой и способностями. Сначала были вычислены «независимые» частоты $n_{i.}n_{.j}/n$. Например, частота $n_{1.}n_{.1}/n$ равна $636 \times 130/1725 = 47{,}030$, а соответствующее ей слагаемое $nD_{11}^2/n_{1.}n_{.1}$ в (33.62) равно

$$(33 - 47{,}930)^2/47{,}930 = 4{,}651.$$

Сумма 24 таких слагаемых оказалась равной $X^2 = 174{,}92$.

Можно вычислить X^2 быстрее, если воспользоваться крайней правой частью равенства (33.62), т. е.

$$X^2 = n\left\{\sum_{i,j}\frac{n_{ij}^2}{n_{i.}n_{.j}} - 1\right\}. \qquad (33.66)$$

Эта операция очень быстро выполняется с помощью счетной машины. Для этого сначала квадрат каждой частоты в i-й строке делится на сумму частот этой строки, а затем полученные частные в j-м столбце делятся на сумму исходных частот этого столбца. Мы должны сначала получить

$$\frac{33^2}{636},\ \frac{48^2}{636},\ \frac{113^2}{636},\ \frac{209^2}{636},\ \frac{194^2}{636},\ \frac{39^2}{636},$$
$$\frac{41^2}{751},\ \frac{100^2}{751},\ \frac{202^2}{751},\ \frac{255^2}{751},\ \frac{138^2}{751},\ \frac{15^2}{751}$$

и еще две аналогичные строки, а затем разделить столбцы на 130, 219 и т. д. Останется лишь вычесть 1 из суммы полученных 24 чисел и умножить результат на 1725 ($=n$), чтобы получить (33.66). Рекомендуем читателю проверить вычисления, используя оба метода.

Получив $X^2 = 174{,}92$, мы находим по формулам (33.63) — (33.65) коэффициенты

$$P = \left(\frac{174{,}92}{1725 + 174{,}92}\right)^{1/2} = 0{,}303,$$

$$T = \left(\frac{174{,}92}{1725\,(3 \times 5)^{1/2}}\right)^{1/2} = 0{,}162, \quad C = \left(\frac{174{,}92}{1725 \times 3}\right)^{1/2} = 0{,}184.$$

Соотношение между этими коэффициентами такое, какого и следовало ожидать, принимая во внимание замечания, сделанные в конце **33.30**: C немного больше T, а P почти в два раза больше, хотя верхняя граница для P равна $(3/4)^{1/2} = 0{,}866$ против $(3/5)^{1/2} = 0{,}880$ для T (и, конечно, 1 для C). Таким образом, верхние границы для P и T в данном случае почти одинаковы, однако P создает впечатление значительно более сильной связи между переменными. В упражнении 33.3 аналогичные результаты получаются для других данных.

Все три коэффициента являются монотонными функциями от X^2, поэтому мы можем проверять независимость, используя непосредственно X^2. В нашем случае $X^2 = 174{,}92$. Так как число степеней свободы равно $(r - 1)(c - 1) = 15$, то полученное значение заведомо превосходит критические значения, обычно используемые на практике. Например, для размера критерия $\alpha = 0{,}001$ критическое значение равно 37,697.

Модели для таблицы $r \times c$

33.31 При рассмотрении мер связи в таблице $r \times c$ необходимо различать исходные модели аналогично тому, как было сделано для таблицы 2×2 в **33.18—23**. С. Рой и Митра (1956) непосредственно обобщили рассмотрение трех типов таблиц (оба набора маргинальных частот фиксированы, один набор фиксирован, ни один набор не фиксирован) на случай таблицы $r \times c$, не встретив при этом никаких новых трудностей. Они показали также (что было сделано нами для таблиц 2×2), что при выполнении гипотезы независимости величина X^2, определенная в (33.62), асимптотически распределена как χ^2 с $(r - 1)(c - 1)$ степенями свободы для всех трех моделей. Интуитивно ясно, что различия между моделями при гипотезе независимости асимптотически исчезают, поскольку все маргинальные частоты, являющиеся случайными величинами, стремятся по вероятности к своим математическим ожиданиям.

Точный анализ таблицы $r \times c$ в духе пункта **33.20** приводит к довольно трудоемким вычислениям. Фримэн и Холтон (1951) рассматривают детально этот метод.

Стандартные ошибки коэффициентов и X^2

33.32 Коэффициенты связи являются монотонными функциями от X^2, поэтому их стандартные ошибки можно получить с помощью (10.14), исходя из стандартной ошибки X^2. Например, стандартная ошибка величины $(X^2)^{1/2}$, к которой (если пренебречь константами) сводятся (33.64) и (33.65), с точностью

до величин порядка n^{-1} равна

$$\mathrm{D}\{(X^2)^{1/2}\} = \left\{\frac{1}{2}(\boldsymbol{X}^2)^{-1/2}\right\}^2 \mathrm{D}X^2 = \frac{\mathrm{D}X^2}{4\boldsymbol{X}^2}, \qquad (33.67)$$

где $\boldsymbol{X}^2$ — генеральное значение X^2, т. е. значение (33.62), вычисленное исходя из частот генеральной совокупности. Очевидно, что первое приближение (33.67) справедливо только при $\boldsymbol{X}^2 \neq 0$, т. е. оно не выполняется, если две категоризованные переменные независимы в генеральной совокупности. Причина этого в том, что выборочная величина X^2 стремится по вероятности к генеральному значению $\boldsymbol{X}^2$, а поскольку $X^2 \geqslant 0$, то распределение X^2 имеет особенность, когда $\boldsymbol{X}^2 = 0$, и его дисперсия имеет порядок n^{-2} (см. аналогичный результат, относящийся к квадрату коэффициента множественной корреляции, в **27.33**). К счастью, мы можем не рассматривать случай независимости, поскольку тогда для критерия можно использовать непосредственно X^2. Если мы хотим оценить ненулевые генеральные коэффициенты $\boldsymbol{T}$ и $\boldsymbol{C}$, выражаемые через $\boldsymbol{X}^2$ формулами, аналогичными (33.64), (33.65), и найти стандартные ошибки оценок, то получим

$$\left.\begin{aligned} \mathrm{D}T &= \frac{1}{4n^2[(r-1)(c-1)]\boldsymbol{T}^2}\mathrm{D}X^2, \qquad \boldsymbol{T} \neq 0, \\ \mathrm{D}C &= \frac{1}{4n^2[\min(r-1,\, c-1)]^2\boldsymbol{C}^2}\mathrm{D}X^2, \quad \boldsymbol{C} \neq 0. \end{aligned}\right\} \qquad (33.68)$$

Для коэффициента Пирсона (33.63) в случае независимости возникает аналогичная трудность, поскольку

$$\left(\frac{\partial P}{\partial (X^2)}\right)^2 = \frac{n^2}{4X^2(n+X^2)^3},$$

так что мы можем написать лишь

$$\mathrm{D}P = \frac{n^2}{4\boldsymbol{X}^2(n+\boldsymbol{X}^2)^3}\mathrm{D}X^2, \quad \text{если} \quad \boldsymbol{X}^2 \neq 0. \qquad (33.69)$$

Выражение (33.69), в отличие от (33.68), не может быть записано в терминах одного только генерального коэффициента $\boldsymbol{P}$.

33.33 Согласно **33.32** для получения требуемых стандартных ошибок нам достаточно вычислить только дисперсию X^2, которая в случае зависимости имеет довольно сложный вид. Для таблицы с фиксированными маргинальными частотами она была вычислена К. Пирсоном (1915), а также (с большей точностью) Янгом и Пирсоном (1915, 1919), которые получили разложение до членов порядка $1/n^3$. Кондо (1929) дал разложения для среднего и дисперсии до порядка $1/n^2$, когда маргинальные частоты случайны, и показал, что в случае независимости дис-

персия имеет порядок $1/n^2$. (Точная дисперсия X^2 в случае независимости вычислена Холдейном (1939) — частный случай его результата сформулирован в упражнении 33.9.) Формулы для дисперсии очень длинны, поэтому мы приведем только первое приближение, полученное К. Пирсоном (1915):

$$\hat{\mathrm{D}}X^2 = 4n\left\{\frac{1}{n}\sum_i\sum_j \frac{(n_{ij}-n_{i.}n_{.j}/n)^2}{(n_{i.}n_{.j}/n)^2} + \frac{X^2}{n} - \left(\frac{X^2}{n}\right)^2\right\}. \quad (33.70)$$

Если подставить (33.70) в (33.68), (33.69) и заменить $\boldsymbol{X^2}$, $\boldsymbol{T}$, $\boldsymbol{C}$ на X^2, T, C, то мы получим требуемые стандартные ошибки.

33.34 Сумма в правой части (33.70) отличается от определения X^2 (33.62) только тем, что знаменатель возведен в квадрат. Если все маргинальные частоты возрастают с возрастанием n, то доминирующим слагаемым в правой части (33.70) будет среднее слагаемое в фигурных скобках, так что асимптотически можно оценивать дисперсию X^2 величиной $4nX^2/n = 4X^2$. Это можно понять и непосредственно. В условиях пункта **30.27** X^2 асимптотически имеет нецентральное распределение χ^2 (в данном случае с $\nu = (r-1)(c-1)$ степенями свободы и параметром нецентральности (30.62), где p_{0i} обозначают «независимые» частоты). Согласно упражнению 24.1 дисперсия такого распределения равна

$$2(\nu + 2\lambda) = 2\left\{(r-1)(c-1) + 2n\sum_i\sum_j \frac{(p_{ij}-p_{i.}p_{.j})^2}{p_{i.}p_{.j}}\right\}.$$

При больших n первым членом в правой части можно пренебречь*), второй же оценивается выражением

$$4n\sum_i\sum_j \frac{(n_{ij}/n - n_{i.}n_{.j}/n^2)^2}{n_{i.}n_{.j}/n^2} = 4X^2.$$

Таким образом, основное слагаемое в стандартной ошибке X^2 порождается асимптотическим нецентральным χ^2-распределением величины X^2. Полезно заметить, что параметр нецентральности λ распределения X^2 оценивается самой величиной X^2, так что использование величины X^2 и ее стандартной ошибки эквивалентно нахождению приближенных границ для λ. Балмер (1958) рассматривает доверительные границы для $\lambda^{1/2}$, естественного параметра «удаленности».

Подстановка $\mathrm{D}X^2 = 4X^2$ в (33.68) дает после упрощений приближения

$$\mathrm{D}T = 1/\{n[(r-1)(c-1)]^{1/2}\}, \quad \mathrm{D}C = 1/\{n\min(r-1)(c-1)\}.$$

*) За исключением случая независимости, когда второй член равен нулю, — это и есть особенность, упомянутая в **33.32** выше.

Другие меры связи

33.35 Вряд ли можно считать, что хотя бы один из коэффициентов (33.63)—(33.65) является удовлетворительной мерой связи. Основной причиной служит то, что эти коэффициенты не имеют простой вероятностной интерпретации (см. Блейлок (1958)). Несколько более легко интерпретируемых мер были предложены Гудмэном и Крускалом (1954, 1959). Мы уже видели в **33.11—14**, что две из предложенных ими мер сводятся в случае таблицы 2×2 к двум общепринятым мерам связи. Для общей таблицы $r \times c$ это не так, и коэффициенты Гудмэна — Крускала не являются функциями от X^2 в отличие от мер, рассмотренных до сих пор. Так, обобщая подход пунктов **33.13—14**, они определяют генеральный коэффициент, равный относительному уменьшению вероятности неправильного предсказания одной переменной при известной второй, а также симметризованный коэффициент, учитывающий предсказуемость в обе стороны. Для упорядоченных классификаций подход пункта **33.12** обобщается аналогичным образом — мы коснемся этого вопроса в связи с рассмотрением упорядоченных таблиц в **33.40** ниже. Гудмэн и Крускал (1963) получили формулы для стандартных ошибок некоторых своих коэффициентов для различных выборочных моделей. См. также Гудмэн (1964b).

Лэнкастер и Хэмдэн (1964), обобщая тетрахорический метод из **26.27—29**, предлагают полихорический метод оценивания коэффициента корреляции для таблицы $r \times c$. Этот метод использует две системы из $(r-1)$ и $(c-1)$ ортонормальных функций (см. **33.44—45**) и требует применения вычислительной машины, но зато он дает значительно лучшие результаты, чем определенный в (33.63) коэффициент P, сводящийся, по существу (см. **33.34**), к приравниванию величины X^2 к своей асимптотической величине в случае двумерного нормального распределения, равной $N\rho^2(1-\rho^2)^{-1}$, и часто дающий неоправданно низкие оценки.

Можно определить статистику отношения правдоподобия, асимптотически эквивалентную X^2, как это делается в **30.5** и упражнении 30.11. Габриэл (1966) показывает, как с помощью такой статистики можно проверять гипотезы, связанные с подмножествами сравниваемых категорий и совокупностей (см. рассмотрение совместных процедур проверки гипотез в **35.54**, **35.63** (том 3)).

Упорядоченные таблицы: ранговые меры связи

33.36 Если существует естественное упорядочение (см. **33.1**) категорий в строках и столбцах таблицы $r \times c$, то мы встречаемся с новой ситуацией, которая не проявлялась в таблице 2×2, потому что в случае двух категорий два возможных порядка категорий могут при любой мере связи только привести к изменению ее знака. В случае же трех или более категорий знание того, что категории упорядочены, дает статистическую информацию, которую можно использовать при измерении

связи. Обычно мы не можем установить какой-либо метрики для категорий; мы знаем, что категории упорядочены, скажем, от более «высокого» к более «низкому» значению исходной переменной, однако мы не можем поставить им в соответствие никаких числовых значений. В таких случаях мы можем применить, возможно несколько неожиданно, ранговую статистику t, рассмотренную в **31.24**. Действительно, можно рассматривать таблицу $r \times c$ с общим числом наблюдений n как способ ранжирования n объектов по двум переменным, для одной из которых можно различать только r рангов, а для другой c рангов. С этой точки зрения маргинальные частоты в таблице суть числа наблюдений, «совпавшие» (см. **31.81**) в различных рангах. Случаю, когда нет ни одного «совпадения», соответствует таблица с маргинальными частотами, равными единице.

33.37 Измерение связи представляется теперь как задача измерения корреляции между двумя ранжировками. Для этого может быть использован любой из коэффициентов t и r_s, определенных в (31.23) и (31.40), однако мы рассмотрим только первый из них.

Здесь возникают некоторые проблемы, связанные с тем, что нас теперь интересуют ранжировки с большим числом совпадений. Во-первых, мы уже не можем определить коэффициент ранговой корреляции с помощью простой системы меток 0—1, как при определении h_{ij} в (31.36), так как теперь имеется три возможности вместо двух. Поэтому мы положим

$$a_{ij} = \begin{cases} +1, & \text{если } x_i < x_j, \\ 0, & \text{если } x_i = x_j, \\ -1, & \text{если } x_i > x_j, \end{cases} \qquad b_{ij} = \begin{cases} +1, & \text{если } y_i < y_j, \\ 0, & \text{если } y_i = y_j, \\ -1, & \text{если } y_i > y_j. \end{cases}$$

Наша мера ранговой корреляции будет основана на сумме

$$S = \sum_{i,\ j} a_{ij} b_{ij}, \quad i,\ j = 1,\ 2,\ \ldots,\ n;\ \ i \neq j. \tag{33.71}$$

Если мы хотим нормировать сумму S так, чтобы она принимала значения в интервале $(-1, +1)$ и достигала крайних значений в случаях полной отрицательной и положительной связанности, удовлетворяя тем самым требованиям пункта **33.5**, то мы имеем несколько возможностей.

(1) Если бы не было ни одного совпадения, то ни одна из величин a_{ij} и b_{ij} не была бы равна 0 и сумма (33.71) изменялась бы от $-n(n-1)$ до $+n(n-1)$ включительно. Мерой связи тогда служило бы отношение $S/\{n(n-1)\}$. Читатель может убедиться, используя определения a_{ij}, b_{ij} и h_{ij}, что это отношение совпадает с t из (31.23). Если некоторые из a_{ij}, b_{ij}

равны нулю, то эта мера, которую мы запишем в виде

$$t_a = \frac{S}{n(n-1)}, \tag{33.72}$$

не может больше достигать ± 1; ее истинные границы изменения зависят от числа нулей.

(2) Если переписать знаменатель (33.72) для случая, когда нет совпадений, в виде $n(n-1) = \left\{ \sum_{i,j} a_{ij}^2 \sum_{i,j} b_{ij}^2 \right\}^{1/2}$ (такая запись показывает, что t является коэффициентом корреляции между двумя наборами меток (см. Даниэлс (1944)), то мы можем определить коэффициент

$$t_b = \frac{\sum_{i,j} a_{ij} b_{ij}}{\left\{ \sum_{i,j} a_{ij}^2 \sum_{i,j} b_{ij}^2 \right\}^{1/2}}. \tag{33.73}$$

Коэффициенты t_a и t_b совпадают, если нет нулевых меток, в противном случае знаменатель (33.73) меньше знаменателя (33.72), так что $t_b > t_a$. Несмотря на это, t_b не может, вообще говоря, достигать ± 1, так как неравенство Коши $\left(\sum a_{ij} b_{ij}\right)^2 \leqslant \leqslant \sum a_{ij}^2 \sum b_{ij}^2$ превращается в равенство, только когда наборы меток a_{ij}, b_{ij} пропорциональны, а для этого в данном случае все наблюдения должны быть сосредоточены на положительной или отрицательной главной диагонали таблицы. Если ни одна из маргинальных частот не равна нулю, то это значит, что t_b может достигать ± 1 только для квадратной таблицы (т. е. таблицы $r \times r$).

(3) Для прямоугольной таблицы $r \times c$ $(r \neq c)$ величина $\left| \sum_{i,j} a_{ij} b_{ij} \right|$ достигает своего максимума, когда все наблюдения лежат в клетках самой длинной диагонали таблицы (т. е. диагонали, содержащей $m = \min(r, c)$ клеток) и насколько возможно поровну распределены между этими клетками. Читатель может проверить, что если n кратно m (что мы можем предположить, поскольку обычно n велико, а m мало), то этот максимум равен $n^2(m-1)/m$. Таким образом, третья мера определяется формулой

$$t_c = \frac{m \sum_{i,j} a_{ij} b_{ij}}{n^2 (m-1)}. \tag{33.74}$$

Коэффициент t_c может достигать ± 1 для любой таблицы $r \times c$, если не считать незначительного эффекта, возникающего, когда n не кратно m. Для больших n, как следует из (33.72) и (33.74), t_c примерно равно $mt_a/(m-1)$.

33.38 Если и маргинальный столбец и маргинальная строка состоят из примерно равных между собой частот, то коэффициенты t_b и t_c отличаются мало. Действительно,

$$\sum_{i,j} a_{ij}^2 = n(n-1) - \sum_{p=1}^{c} n_{.p}(n_{.p}-1) = n^2 - \sum_{p=1}^{c} n_{.p}^2$$

и, аналогично, $\sum_{i,j} b_{ij}^2 = n^2 - \sum_{p=1}^{r} n_{p.}^2$. Таким образом, знаменатель t_b равен

$$\left(\sum a_{ij}^2 \sum b_{ij}^2\right)^{1/2} = \left\{\left(n^2 - \sum_{p=1}^{c} n_{.p}^2\right)\left(n^2 - \sum_{p=1}^{r} n_{p.}^2\right)\right\}^{1/2} = \tag{33.75}$$

$$= n^2\left\{\left(1 - \frac{1}{n^2}\sum_{p=1}^{c} n_{.p}^2\right)\left(1 - \frac{1}{n^2}\sum_{p=1}^{r} n_{p.}^2\right)\right\}, \tag{33.76}$$

тогда как знаменатель t_c равен

$$n^2\,\frac{m-1}{m} = n^2\left(1 - \frac{1}{m}\right). \tag{33.77}$$

Если все частоты $n_{.p}$ в маргинальном столбце равны, а также равны все частоты $n_{p.}$ в маргинальной строке, то (33.76) приближенно равно

$$n^2\left\{\left(1 - \frac{1}{c}\right)\left(1 - \frac{1}{r}\right)\right\}^{1/2}. \tag{33.78}$$

(33.78) совпадает с (33.77), если таблица квадратная ($r = c = m$), в противном же случае (33.78) больше и, следовательно, t_b меньше. С другой стороны, когда маргинальные частоты не равны точно, суммы квадратов возрастают и, следовательно, знаменатель t_b (33.76) убывает.

Следующий пример (см. также Кендалл (1962)) иллюстрирует вычисление коэффициентов.

Пример 33.8

В представленной ниже таблице нас интересует связь между дальностью зрения правым и левым глазом.

Для вычисления числителя всех коэффициентов t каждая частота умножается на все частоты к юго-востоку от нее со знаком плюс и на все частоты к юго-западу со знаком минус. Клетки той же строки и того же столбца не учитываются. (Нет необходимости применять процедуру к последней строке таблицы, поскольку ниже нее ничего нет.) Выражение $\sum_{i,j} a_{ij}b_{ij}$ равно

Таблица 33.3

3242 мужчин 30—39 лет — служащих Королевских артиллерийских заводов Великобритании (1943—1946): дальность зрения невооруженным глазом

Правый глаз \ Левый глаз	Высшая степень	Вторая степень	Третья степень	Низшая степень	Сумма
Высшая степень	821	112	85	35	1053
Вторая степень	116	494	145	27	782
Третья степень	72	151	583	87	893
Низшая степень	43	34	106	331	514
Сумма	1052	791	919	480	3242

удвоенной сумме таких слагаемых, потому что мы можем иметь как $i < j$, так и $i > j$. Для нашей таблицы получаем

$$821\,(494 + 145 + 27 + 151 + 583 + 87 + 34 + 106 + 331) + \\ + 112\,(145 + 27 + 583 + 87 + 106 + 331 - 116 - 72 - 43)$$

и т. д. При переходе к нижней части таблицы число слагаемых в скобках убывает. Читателю предлагается проверить, что после сложения слагаемых в скобках мы получим

$$821 \cdot 1958 + 112 \cdot 1048 + 85 \cdot (-465) + 35 \cdot (-1744) + \\ + 116 \cdot 1292 + 494 \cdot 992 + 145 \cdot 118 + 27 \cdot (-989) + 72 \cdot 471 + \\ + 151 \cdot 394 + 583 \cdot 254 + 87 \cdot (-183) = 2\,480\,223.$$

Таким образом, числитель равен $\sum a_{ij}b_{ij} = 4\,960\,446$. Пользуясь формулой (33.75), находим знаменатель t_b:

$$[\{3242^2 - (1052^2 + 791^2 + 919^2 + 480^2)\} \times \\ \times \{3242^2 - (1053^2 + 782^2 + 893^2 + 514^2)\}]^{1/2} = \\ = [7\,728\,586 \times 7\,703\,218]^{1/2}.$$

Следовательно,

$$t_b = \frac{4\,960\,446}{[7\,728\,586 \times 7\,703\,218]^{1/2}} = 0{,}643.$$

С другой стороны, по формуле (33.74) находим

$$t_c = \frac{4 \times 4\,960\,446}{3242^2 \times 3} = 0{,}629.$$

Итак, в данном случае, когда наибольшие маргинальные частоты в каждом из двух наборов примерно в два раза больше наименьших, t_b лишь незначительно больше t_c. Аналогичный результат получается в упражнении 33.10, где максимальные

маргинальные частоты примерно в три раза больше минимальных.

33.39 Если не учитывать достижимости ± 1, рассмотренной в **33.37**, основное отличие t_b от t_c состоит в том, что для стандартной ошибки t_c можно найти верхнюю границу (см. (33.81) ниже) в случае выборки из n наблюдений с нефиксированными маргинальными частотами, тогда как коэффициент t_b в этой ситуации является отношением двух случайных величин и его стандартная ошибка неизвестна. Если маргинальные частоты фиксированы, то t_b уже не будет отношением двух случайных величин, однако его распределение изучено только в предположении независимости двух категоризованных переменных в таблице. Конечно, если надо проверить независимость, то можно ограничиться общим числителем $\sum a_{ij}b_{ij}$. Детали, касающиеся этого критерия, даются Кендаллом (1962). Стьюарт (1953) показал, что верхние границы дисперсии t_c могут быть использованы для проверки гипотез о разности двух значений, полученных для разных таблиц. Поскольку это довольно очевидно, то мы не рассматриваем данный вопрос.

33.40 Гудмэн и Крускал (1954) предложили меру связи для упорядоченных таблиц, тесно связанную с коэффициентами t, которые мы рассмотрели. Эта мера имеет вид

$$G = \frac{\sum a_{ij}b_{ij}}{n^2 - \sum_{p=1}^{c} n_{\cdot p}^2 - \sum_{p=1}^{r} n_{p\cdot}^2 + \sum_{p=1}^{r}\sum_{q=1}^{c} n_{pq}^2}, \tag{33.79}$$

отличаясь от t только знаменателем.

Знаменатель t_b в (33.75) тождественно равен

$$\left\{\left[n^2 - \frac{1}{2}\left(\sum_p n_{\cdot p}^2 + \sum_p n_{p\cdot}^2\right)\right]^2 - \frac{1}{4}\left(\sum_p n_{\cdot p}^2 - \sum_p n_{p\cdot}^2\right)^2\right\}^{1/2},$$

т. е. очень близок к

$$n^2 - \frac{1}{2}\left(\sum_p n_{\cdot p}^2 + \sum_p n_{p\cdot}^2\right).$$

Сравнивая эту величину со знаменателем G, мы видим, что последний практически всегда меньше. Легко видеть, что G может достигать своих пределов ± 1, если все наблюдения лежат на самой длинной диагонали таблицы. Таким образом, G скорее ближе к t_c. Гудмэн и Крускал (1963) дали стандартную ошибку G и метод ее вычисления, а также получили простую верхнюю границу

$$\mathrm{D}G \leqslant \frac{2}{n}\,\frac{(1-G)^2}{D_G/n^2}, \tag{33.80}$$

где D_G — знаменатель G в (33.79). Эту границу можно сравнить с верхней границей для дисперсии t_c:

$$\mathrm{D}t_c \leqslant \frac{2}{n}\left\{\left(\frac{m}{m-1}\right)^2 - t_c^2\right\}. \qquad (33.81)$$

Гудмэн и Крускал (1963) показали на двух примерах, что G имеет тенденцию быть больше t_c, однако стандартная ошибка G значительно меньше. Детали приведены в упражнении 33.11. Если бы это было доказано в общем случае, то этот факт, наряду с непосредственной интерпретируемостью G в терминах отношений порядка в случайной выборке (G дает вероятность того, что иксы расположены в том же порядке, что и игреки, минус вероятность расположения в противоположном порядке при условии, что нет совпадений, — см. **33.12** для случая 2×2), сделал бы G стандартной мерой связи в случае упорядоченных таблиц.

Эксперименты Розенталя (1966) для малых выборок имеют целью проверить приложимость асимптотической теории к G.

Методы, использующие заранее заданные системы меток

33.41 Возвращаясь к общему рассмотрению таблиц $r \times c$, мы рассмотрим возможность введения метрики для категорий таблицы. Приписывая числовые метки категориям каждой переменной, мы сведем задачу измерения зависимости и взаимозависимости к анализу обычной (группированной) двумерной таблицы, подробно изученной в главе 26. Так мы сможем обычным способом вычислить коэффициенты корреляции и регрессии. Трудность заключается в выборе подходящей системы меток. Мы уже обсуждали этот вопрос с точки зрения ранговых критериев в **31.21—24**, где было показано, что для различных систем «условных чисел» (т. е. «меток» в принятой терминологии) получаются различные критерии. В рассматриваемом случае трудности возрастают, так как строится мера, а не просто критерий независимости.

Простейшая система меток использует наборы натуральных чисел $1, 2, \ldots, r$ и $1, 2, \ldots, c$ для категорий, соответствующих строкам и столбцам. Можно использовать также наборы нормальных меток $E(s, r)$ и $E(s, c)$, рассмотренные в **31.39**. Пример 33.9 иллюстрирует эти процедуры.

Пример 33.9

Вычислим для данных примера 33.8 коэффициенты корреляции, используя системы меток пункта **33.41**. При использовании натуральных чисел мы припишем метки 1, 2, 3, 4 категориям от «высшей» к «низшей» для левого глаза (x) и анало-

гично (поскольку таблица оказывается здесь квадратной) для правого глаза (y). Мы получаем

$$\sum x = (1052 \cdot 1) + (791 \cdot 2) + (919 \cdot 3) + (480 \cdot 4) = 7311,$$
$$\sum y = (1053 \cdot 1) + (782 \cdot 2) + (893 \cdot 3) + (514 \cdot 4) = 7352,$$
$$\sum x^2 = (1052 \cdot 1^2) + \ldots = 20\,167,$$
$$\sum y^2 = (1053 \cdot 1^2) + \ldots = 20\,442,$$
$$\sum xy = (821 \cdot 1 \cdot 1) + (112 \cdot 1 \cdot 2) + \ldots = 19\,159.$$

Таким образом, коэффициент корреляции для системы меток, использующей натуральные числа, равен

$$r_1 = \frac{19\,159 - 7311 \cdot 7352/3242}{[\{20\,167 - (7311)^2/3242\}\{20\,442 - (7352)^2/3242\}]^{1/2}} = \frac{2579}{(3677 \cdot 3772)^{1/2}} = 0{,}69.$$

Полученный результат не очень отличается от значений ранговых мер $t_b = 0{,}64$ и $t_c = 0{,}63$, найденных в примере 33.8, чего, впрочем, следовало ожидать, принимая во внимание, что система меток, использующая натуральные числа, тесно связана, как мы видели в главе 31, с коэффициентами ранговой корреляции.

Предположим теперь, что мы используем нормальные метки

$$E(1,\ 4) = -1{,}029, \quad E(2,\ 4) = -0{,}297,$$
$$E(3,\ 4) = +0{,}297, \quad E(4,\ 4) = +1{,}029,$$

полученные из *Biometrika Tables* (таблица 28). Находим

$$\sum x = 1{,}029\,(480 - 1052) + 0{,}297\,(919 - 791) = -550{,}6,$$
$$\sum y = 1{,}029\,(514 - 1053) + 0{,}297\,(893 - 782) = -521{,}7,$$
$$\sum x^2 = (1{,}029)^2\,(480 + 1052) + (0{,}297)^2\,(919 + 791) = 1773,$$
$$\sum y^2 = (1{,}029)^2\,(514 + 1053) + (0{,}297)^2\,(893 + 782) = 1807,$$
$$\sum xy = (1{,}029)^2\,(821 + 331 - 35 - 43) + (0{,}297)^2\,(494 + 583 - 145 - 151) + (1{,}029)\,(0{,}297)\,(116 + 112 + 106 + 87 - 72 - 34 - 85 - 27) = 1268,$$

откуда получаем коэффициент корреляции для нормальных меток

$$r_2 = \frac{1268 - (550{,}6)\,(521{,}7)/3242}{[\{1773 - (550{,}6)^2/3242\}\{1807 - (521{,}7)^2/3242\}]^{1/2}} = \frac{1179}{(1680 \cdot 1723)^{1/2}} = 0{,}69.$$

То же самое значение, с точностью до двух десятичных знаков, было получено при использовании натуральных чисел. Принимая во внимание значительное усложнение вычислений, связанное с нормальными метками, вряд ли стоит прибегать к ним, по крайней мере при таком малом числе категорий, как 4.

33.42 Если уж вводить нормальную метрику в таблице $r \times c$, то разумнее было бы приписать категориям числовые значения, которые соответствовали бы наблюденным частотам при выборе из нормального распределения. Так, в примере 33.9 следовало бы вычислить «точки разбиения» стандартного нормального распределения, соответствующие частотам $\frac{1052}{3242}$, $\frac{791}{3242}$, $\frac{919}{3242}$, $\frac{480}{3242}$, и взять в качестве меток категорий для левого глаза средние четырех частей нормального распределения.

В данный момент мы не будем проводить вычислений, однако ясно, что полученные метки будут отличаться от грубых нормальных меток, использованных в примере 33.9. Мы вернемся к этой системе меток в **33.50** ниже.

На этом мы прекращаем изучение методов, в которых используются заранее выбранные системы меток, так как ясно, что, добавляя «информацию» к таблице таким путем, мы делаем некоторые предположения о распределениях, что может привести к ошибкам в случае неправильности этих предположений. Вообще же этого следовало бы избегать, предпочитая использование ранговых методов, рассмотренных в **33.36—40**. Иэйтс (1948) первым предложил систему меток с натуральными числами пункта **33.41**, а Э. Уилльямс (1952) дал общий обзор методов, использующих метки.

Анализ таблиц с помощью наименьших квадратов

33.43 Если изучается зависимость одной категоризации (скажем, соответствующей строкам) от другой, то можно избежать введения метрики для последней. Определим для каждого наблюдения набор меток x_i $(i = 1, 2, \ldots, c-1)$ следующим образом: $x_i = +1$, если наблюдение находится в i-м столбце, и $x_i = 0$ в противном случае. Совокупность $c-1$ меток определяет для каждого наблюдения его столбец, при этом принадлежности наблюдения к c-му столбцу соответствует равенство нулю всех $c-1$ меток. Далее, если категоризация, соответствующая строкам, имеет (или ей поставили в соответствие) метрику, то можно применить обычный метод наименьших квадратов главы 19 для анализа зависимости этой метрики от меток x. При этом каждому x_i будет соответствовать параметр θ_i. Никаких новых теоретических трудностей при этом не возникает.

Выбор «оптимальных» меток: канонический анализ

33.44 Можно, однако, подойти к задаче введения меток для категорий таблицы $r \times c$ (не обязательно упорядоченной) с совсем другой точки зрения. Мы можем поставить вопрос: какие метки надо сопоставить категориям, чтобы коэффициент корреляции между двумя переменными был максимальным? Оказывается, причем довольно неожиданно, что такие «оптимальные» метки тесно связаны с преобразованием частот таблицы в частоты двумерного нормального распределения. Сначала мы докажем одну теорему (Лэнкастер (1957)) для несгруппированных наблюдений.

Пусть x и y имеют двумерное нормальное распределение с коэффициентом корреляции ρ. Пусть $x' = x'(x)$ и $y' = y'(y)$ — новые переменные, зависящие соответственно только от x и только от y, для которых $\mathbf{M}\{(x')^2\}$ и $\mathbf{M}\{(y')^2\}$ конечны. Тогда справедливо представление

$$x' = a_0 + a_1 H_1(x) + a_2 H_2(x) + \ldots, \tag{33.82}$$

где H_r — полиномы Чебышева — Эрмита, определенные формулами (6.21) и нормированные условием

$$\int_{-\infty}^{\infty} H_r^2(x)\,\alpha(x)\,dx = 1.$$

Ряд $\sum_{i=1}^{\infty} a_i^2$ сходится. Так как коэффициент корреляции не меняется при изменениях положения и масштаба, то можно положить $a_0 = 0$ и написать $x' = \sum_{i=1}^{\infty} a_i H_i$, $\sum_{i=1}^{\infty} a_i^2 = 1$, и, аналогично, $y' = \sum_{i=1}^{\infty} b_i H_i$, $\sum_{i=1}^{\infty} b_i^2 = 1$. В соответствии с **6.14** полином $H_r(x)$ равен коэффициенту при $t^r/r!$ в разложении $\exp\left(tx - \frac{1}{2}t^2\right)$. Поскольку математическое ожидание $\exp\left(tx - \frac{1}{2}t^2 + uy - \frac{1}{2}u^2\right)$ равно $\exp(\rho tu)$, то мы получим

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_r(x)\,H_s(y)\,f\,dx\,dy = \begin{cases} \rho^r, & r = s, \\ 0, & r \neq s, \end{cases} \tag{33.83}$$

где f — плотность двумерного нормального распределения. Дисперсии x' и y' в силу ортогональности H_r равны единице, так что коэффициент корреляции равен

$$\operatorname{cov}(x', y') = \sum_{i=1}^{\infty} a_i b_i \rho^i. \tag{33.84}$$

Он меньше, чем $|\rho|$, за исключением случая $a_1^2 = b_1^2 = 1$. Остальные a и b при этом обращаются в нуль. Таким образом, максимум корреляции между x' и y' равен $|\rho|$, и теорема Лэнкастера формулируется следующим образом: если двумерное распределение (x, y) может быть получено из двумерного нормального распределения с помощью преобразований x и y (каждого отдельно), то корреляция преобразованного распределения не может по абсолютной величине превзойти корреляции ρ двумерного нормального распределения.

33.45 Предположим, что мы ищем теперь вторую такую пару преобразований x и y, скажем x'' и y''. Если потребовать, чтобы x'' и y'' были нормированы и некоррелированы с первой парой (x', y'), то из ортогональности разложений Чебышева — Эрмита $x'' = \sum_{i=1}^{\infty} c_i H_i(x)$, $y'' = \sum_{i=1}^{\infty} d_i H_i(y)$ будет следовать, что $c_1 = d_1 = 0$. Таким образом, мы получим $x'' = \sum_{i=2}^{\infty} c_i H_i$, $y'' = \sum_{i=2}^{\infty} d_i H_i$ и, аналогично (33.84),

$$\operatorname{cov}(x'', y'') = \sum_{i=2}^{\infty} c_i d_i \rho^i. \tag{33.85}$$

(33.85) достигает максимума по абсолютной величине, только когда $c_2^2 = d_2^2 = 1$, а все другие c_r, d_r равны 0. В этом случае корреляция равна ρ^2. Можно перейти к следующим парам переменных (x''', y'''), $(x^{(\mathrm{IV})}, y^{(\mathrm{IV})})$ и т. д., получая максимальные корреляции $|\rho|^3$, ρ^4 и т. д.

Преобразованные пары переменных называются *каноническими переменными*. Мы показали, что r-я пара канонических переменных имеет *каноническую корреляцию* $|\rho|^r$. Из нашего доказательства ясно, что сами канонические переменные являются просто формами Чебышева — Эрмита от двумерных нормальных переменных (x, y), т. е.

$$(x^{(r)}, y^{(r)}) = (H_r(x), H_r(y)). \tag{33.86}$$

Лэнкастер (1958) развивает дальше этот подход.

33.46 Результаты пунктов **33.44—45** применимы к несгруппированным значениям двумерного распределения. На практике, имея выборку в виде упорядоченной таблицы $r \times c$, не представляет труда найти такие преобразования каждой из переменных, которые привели бы к одномерным нормальным маргинальным распределениям, — именно это предлагалось в **33.42**. Однако лишь по счастливой случайности эти преобразования могут дать двумерные нормальные частоты в самой таблице. Все

же теоретический смысл полученного результата ясен: стараясь найти системы меток для каждой из категоризованных переменных, максимизирующие их корреляцию, мы, по существу, стараемся получить двумерное нормальное распределение, оперируя только маргинальными частотами таблицы.

33.47 Предположим, что мы сопоставили категориям таблицы $r \times c$ метки x_i, y_j $(i = 1, 2, \ldots, r;\ j = 1, 2, \ldots, c)$. Не теряя общности, можно считать их нормированными (нулевое среднее, единичная дисперсия). Мы имеем тогда

$$\mathsf{D}x = \sum_i n_{i.} x_i^2/n = \mathsf{D}y = \sum_j n_{.j} y_j^2/n = 1, \tag{33.87}$$

$$\operatorname{cov}(x, y) = \operatorname{corr}(x, y) = \sum_i \sum_j n_{ij} x_i y_j/n. \tag{33.88}$$

Требуется максимизировать (33.88) по x и y при условии (33.87). Вводя неопределенные множители Лагранжа, λ и μ, мы получим систему уравнений

$$\sum_j n_{ij} y_j - \lambda n_{i.} x_i = 0, \quad i = 1, 2, \ldots, r, \tag{33.89}$$

$$\sum_i n_{ij} x_i - \mu n_{.j} y_j = 0, \quad j = 1, 2, \ldots, c. \tag{33.90}$$

Умножая (33.89) на x_i и суммируя по i, получаем $\lambda = R$, где R — корреляция, которую мы ищем. Аналогично находим, что $\mu = R$, и, следовательно, в силу (33.89), (33.90) приходим к системе

$$\left.\begin{aligned} \sum_j n_{ij} y_j &= R n_{i.} x_i, \\ \sum_i n_{ij} x_i &= R n_{.j} y_j. \end{aligned}\right\} \tag{33.91}$$

Исключение x и y приводит к детерминантному уравнению, которое символически может быть записано в виде

$$\begin{vmatrix} R n_{i.} & n_{ij} \\ n_{ij} & R n_{.j} \end{vmatrix} = 0. \tag{33.92}$$

Мы вернемся к этому уравнению при изучении канонических корреляций в томе 3. Здесь достаточно лишь отметить, что R можно выразить через частоты таблицы. Действительно, (33.92) есть уравнение относительно R^2 с несколькими корнями. В общем случае имеется $m = \min(r, c)$ ненулевых корней, один из которых всегда равен единице и который для нас не интересен. Остальные $m - 1$ корней называются каноническими корреляциями. Число ненулевых канонических корреляций не может

быть больше $m-1$, поскольку ранг матрицы частот $\{n_{ij}\}$ не превышает m. Нам нужен наибольший корень R_1. Остальные с точностью до выборочных флуктуаций являются, как мы видели в **33.45**, степенями наибольшего корня.

33.48 Из (33.92) следует, что если канонические корреляции (корни (33.92)) равны $R_1, R_2, \ldots, R_{m-1}$, то

$$n_{ij} = \frac{n_{i.}n_{.j}}{n}\left\{1 + \sum_{s=1}^{m-1} R_s x_s y_s\right\}. \tag{33.93}$$

В пределе, при стремлении m к бесконечности и измельчении категорий таблицы $r \times c$, (33.93) в силу **33.44—45** сводится к тетрахорическому ряду

$$f = (2\pi)^{-1}\exp\left\{-\tfrac{1}{2}(x^2+y^2)\right\}\left\{1 + \sum_{j=1}^{\infty} H_j(x)\,H_j(y)\,\rho^j\right\}, \tag{33.94}$$

где f — плотность двумерного нормального распределения. Формула (33.94) является просто другой записью формулы (26.66), которая отличается только множителем $j!$ в знаменателе, поскольку в (26.66) H_j не были нормированы.

33.49 Вычислив с помощью (33.92) наибольшую каноническую корреляцию R_1, можно сразу найти «оптимальные» наборы меток, приводящие к этой корреляции. Вероятно, это легче всего сделать, обращаясь к (33.91). Если умножить в (33.91) второе уравнение на $n_{pj}/\{(n_{p.}n_{i.})^{1/2}\,n_{.j}\}$ и просуммировать по j, то оно примет вид

$$\sum_i \left(\sum_j n_{ij}x_i\right)\frac{n_{pj}}{(n_{p.}n_{i.})^{1/2}\,n_{.j}} = \frac{R\sum_j n_{pj}y}{(n_{p.}n_{i.})^{1/2}} = \left(\frac{n_{p.}}{n_{i.}}\right)^{1/2} R^2 x_p. \tag{33.95}$$

Это можно переписать в виде

$$\sum_i (x_i n_{i.}^{1/2})\sum_j \frac{n_{ij}n_{pj}}{(n_{p.}n_{i.})^{1/2}\,n_{.j}} = R^2(x_p n_{p.}^{1/2}), \tag{33.96}$$

откуда ясно, что квадраты канонических корреляций служат характеристическими корнями $(r \times r)$-матрицы $\boldsymbol{NN'}$, где $\boldsymbol{N}$ — матрица $(r \times c)$ с элементами $n_{ij}/(n_{i.}n_{.j})^{1/2}$. Действительно, (33.96) в матричной форме имеет вид

$$\boldsymbol{NN'u} = R^2\boldsymbol{u}, \tag{33.97}$$

где $\boldsymbol{u}$ — вектор-столбец размерности r с компонентами $x_i n_{i.}^{1/2}$.

Так как ранг $\boldsymbol{N}$, а следовательно и $\boldsymbol{NN'}$, не превосходит $m = \min(r, c)$, то $\boldsymbol{NN'}$ имеет в общем случае m ненулевых корней, как и утверждалось в **33.47**. Легко проверить, что $R^2 = 1$ всегда является корнем. Соответствующий собственный вектор $\boldsymbol{u}$ имеет

компоненты $(n_{p\cdot})^{1/2}$, т. е. для этого корня $x_p \equiv 1$. Оставляя в стороне этот корень, не имеющий отношения к задаче измерения связи, мы рассмотрим наибольший характеристический корень R_1^2, определяемый из (33.97). Соответствующий ему собственный вектор $\boldsymbol{u}_1$ дает набор меток для первой канонической переменной x. Аналогично, метки для первой канонической переменной y даются собственным вектором $\boldsymbol{v}_1$, получаемым из соотношения

$$\boldsymbol{N'Nv} = R^2\boldsymbol{v}, \tag{33.98}$$

где $\boldsymbol{v}$ — вектор-столбец размерности c с элементами $y_j n_{\cdot j}^{1/2}$. Ненулевые характеристические корни матрицы $\boldsymbol{N'N}$, конечно, совпадают с корнями $\boldsymbol{NN'}$, т. е. равны квадратам канонических корреляций. Нет необходимости, однако, решать как (33.97), так и (33.98), достаточно решить одну из этих систем (естественно выбрать ту, которая имеет меньший порядок, например $\boldsymbol{NN'}$, если $r \leqslant c$), а затем найти другой набор меток с помощью системы (33.91), которую мы перепишем в виде

$$\left.\begin{aligned} y_j &= \frac{1}{Rn_{\cdot j}} \sum_i n_{ij} x_i, \\ x_i &= \frac{1}{Rn_{i\cdot}} \sum_j n_{ij} y_j. \end{aligned}\right\} \tag{33.99}$$

Пример 33.10

Применим канонический анализ к данным примера 33.8. Перепишем сначала таблицу, заменив маргинальные частоты их квадратными корнями:

821	112	85	35	32,4499615
116	494	145	27	27,9642629
72	151	583	87	29,8831056
43	34	106	331	22,6715681
32,4345495	28,1247222	30,3150128	21,9089023	

Построим теперь матрицу $\boldsymbol{N}$, деля n_{ij} на произведение соответствующих маргинальных квадратных корней, например, $821/(32{,}4345495 \times 32{,}4499615)$. Получим

$$\boldsymbol{N} = \begin{Vmatrix} 0{,}780047593 & 0{,}122720070 & 0{,}086406614 & 0{,}049230386 \\ 0{,}127892990 & 0{,}628109454 & 0{,}171043616 & 0{,}044069667 \\ 0{,}074284618 & 0{,}179664791 & 0{,}643554112 & 0{,}132884065 \\ 0{,}058476184 & 0{,}053322330 & 0{,}154229180 & 0{,}666385924 \end{Vmatrix}. \tag{33.100}$$

В данном случае $r = c$, так что безразлично, работаем мы с $\boldsymbol{NN'}$ или с $\boldsymbol{N'N}$. Мы найдем

$$\boldsymbol{NN'} = \begin{Vmatrix} 0{,}633424197 & 0{,}193793122 & 0{,}142143279 & 0{,}098290784 \\ 0{,}193793122 & 0{,}442076157 & 0{,}238281615 & 0{,}096718276 \\ 0{,}142143279 & 0{,}238281615 & 0{,}469617711 & 0{,}201730920 \\ 0{,}098290784 & 0{,}096718276 & 0{,}201730920 & 0{,}474119575 \end{Vmatrix}. \qquad (33.101)$$

Вспомним, что сумма характеристических корней равна следу матрицы, так что если след матрицы ненамного превышает 1, то наибольшая каноническая корреляция должна быть мала. Это служит хорошей предварительной проверкой. Здесь след больше 2, так что R_1^2 вполне может быть порядка 1.

Найдем теперь характеристические корни. Для этого надо решить характеристическое уравнение

$$|\boldsymbol{NN'} - \lambda \boldsymbol{I}| = 0. \qquad (33.102)$$

Вычитая λ из каждого диагонального элемента $\boldsymbol{NN'}$ и расписывая определитель получившейся матрицы, получаем уравнение четвертой степени

$$\lambda^4 - 2{,}01923764\lambda^3 + 1{,}343416989\lambda^2 - \\ - 0{,}355747132\,\lambda + 0{,}031567594 = 0.$$

Поскольку один корень этого уравнения равен единице, в левой части можно выделить множитель $(\lambda - 1)$:

$$(\lambda - 1)(\lambda^3 - 1{,}019237640\lambda^2 + 0{,}324179349\lambda - 0{,}031567783) = 0.$$

Мы пришли к кубическому уравнению, которое решается стандартными методами. Его корни дают квадраты канонических корреляций

$$\left.\begin{aligned} R_1^2 &= \lambda_1 = 0{,}48516, \\ R_2^2 &= \lambda_2 = 0{,}34604, \\ R_3^2 &= \lambda_3 = 0{,}18803. \end{aligned}\right\} \qquad (33.103)$$

Заметим, что значение $R_1 = 0{,}697$ ненамного больше значения 0,69, полученного при использовании натуральных чисел и нормальных меток в примере 33.9.

Найдем теперь собственные векторы, соответствующие λ_1. Сначала решим систему уравнений для получения компонент $\boldsymbol{u}_1$:

$$\boldsymbol{NN'u}_1 = 0{,}48516\boldsymbol{u}_1.$$

Разделив затем компоненты u_i вектора $\boldsymbol{u}_1$ на $n_{i\cdot}^{1/2}$, мы найдем канонический набор меток для категорий, соответствующих строкам:

$$\boldsymbol{x} = \begin{Bmatrix} u_1/n_{1\cdot}^{1/2} \\ u_2/n_{2\cdot}^{1/2} \\ u_3/n_{3\cdot}^{1/2} \\ u_4/n_{4\cdot}^{1/2} \end{Bmatrix} = \begin{pmatrix} -1{,}307 \\ +0{,}021 \\ +0{,}739 \\ +1{,}362 \end{pmatrix}. \tag{33.104}$$

Канонические метки для y получаются с помощью (33.99). Так, например,

$$y_1 = \{821 \cdot (-1{,}307) + 116 \cdot 0{,}021 + 72 \cdot 0{,}739 + \\ + 43 \cdot 1{,}362\} / \{1052 \cdot (0{,}48516)^{1/2}\}.$$

Отсюда находим набор меток

$$\boldsymbol{y} = \begin{Bmatrix} -1{,}309 \\ +0{,}040 \\ +0{,}730 \\ +1{,}406 \end{Bmatrix}. \tag{33.105}$$

Наборы меток (33.104) и (33.105), взвешенные в соответствии с частотами по строкам или столбцам, дают нулевые средние, однако их дисперсии не равны единице, так как характеристические векторы определяются с точностью до произвольной масштабной константы.

33.50 Вспомним теперь интерпретацию теоремы Лэнкастера в **33.46**: выбор меток, которые приводят к максимальной корреляции между категоризованными переменными, по существу эквивалентен преобразованию маргинальных частот таблицы к одномерной нормальности с намерением получить двумерную нормальность в таблице. Интересно в связи с этим использовать для обработки данных примера 33.10 нормальную систему меток, описанную в **33.42**. Мы тогда сможем увидеть, насколько хорошо согласуются метки, полученные для категорий, с каноническими метками в примере 33.10.

Пример 33.11

В примере 33.8 мы имеем следующие две совокупности относительных маргинальных частот (в скобках даны соответствующие интервалы изменения стандартной нормальной

случайной величины, полученные из *Biometrika Tables*):

$p_i=\left(\frac{n_{i\cdot}}{n}\right)$	Соответствующий интервал	$p_i=\left(\frac{n_{\cdot i}}{n}\right)$	Соответствующий интервал
0,3245	$(-\infty, -0{,}4551)$	0,3248	$(-\infty, -0{,}4543)$
0,2440	$(-0{,}4551, +0{,}1726)$	0,2412	$(-0{,}4543, +0{,}1662)$
0,2835	$(+0{,}1726, +1{,}0450)$	0,2755	$(+0{,}1662, +1{,}0006)$
0,1480	$(+1{,}0450, +\infty)$	0,1585	$(+1{,}0006, +\infty)$
1,0000		1,0000	

(33.106)

Среднее значение стандартной нормальной случайной величины в интервале (a_i, b_i), имеющем вероятность p_i, равно

$$\mu_i = \frac{1}{p_i}\int_{a_i}^{b_i} (2\pi)^{-1/2} te^{-t^2/2}\,dt \propto \frac{1}{p_i}\left(e^{-a_i^2/2} - e^{-b_i^2/2}\right). \quad (33.107)$$

Мы пренебрегаем множителем $(2\pi)^{-1/2}$, поскольку нам нужна только корреляция, а она не зависит от масштабного множителя. С помощью четырехзначных таблиц логарифмов находим следующие значения меток:

Метки для строк	Метки для столбцов
−2,778	−2,777
−0,343	−0,350
+1,432	+1,380
+3,914	+3,825

В отличие от примера 33.10, полученные значения не будут точно нормированными, даже если учесть отброшенный множитель $(2\pi)^{-1/2}$, поскольку в результате усреднения по интервалам стандартного нормального распределения возникает эффект группировки. Мы получаем следующие значения для сумм и сумм квадратов (взвешенных, конечно, в соответствии с маргинальными частотами):

	Для строк	Для столбцов
Сумма	+98	−94
Сумма квадратов	17923	16984
Среднее	0,030	−0,029
Стандартное отклонение	2,351	2,289

Вычитание среднего и деление на стандартное отклонение приводит к следующим значениям меток:

Нормированные метки для строк	Нормированные метки для столбцов
−1,194	−1,205
−0,159	−0,140
+0,596	+0,616
+1,652	+1,684

(33.108)

Значения (33.108) довольно плохо согласуются с (33.104), (33.105).

Таким образом, метод, использованный в этом примере, дает лишь грубое приближение к канонической системе меток.

Разбиение X^2: канонические компоненты

33.51 Канонические корреляции, рассмотренные в **33.44—49**, тесно связаны с X^2-статистикой (33.62). Рассмотрим снова матрицу $\boldsymbol{NN'}$, определенную в **33.49**. Ее диагональные элементы имеют вид

$$(\boldsymbol{NN'})_{ii} = \sum_j \frac{n_{ij}^2}{n_{i.}n_{.j}},$$

так что в силу (33.62) получаем

$$\operatorname{tr} \boldsymbol{NN'} = \sum_i \sum_j \frac{n_{ij}^2}{n_{i.}n_{.j}} = \frac{X^2}{n} + 1. \tag{33.109}$$

Вспоминая, что след матрицы равен сумме характеристических корней и что характеристические корни $\boldsymbol{NN'}$ равны $1, R_1^2, \ldots, \ldots, R_{m-1}^2$, получаем в соответствии с (33.109)

$$X^2 = n(R_1^2 + R_2^2 + \ldots + R_{m-1}^2). \tag{33.110}$$

Мы представили, таким образом, X^2 в виде умноженной на n суммы компонент, равных квадратам канонических корреляций.

Хотелось бы предположить, что компоненты в (33.110) сами являются при гипотезе независимости асимптотически независимыми случайными величинами, распределенными по закону χ^2 с числом степеней свободы, равным для nR_s^2

$$(r-s)(c-s) - (r-(s+1))(c-(s+1)) = r + c - 2s - 1.$$

Однако Лэнкастер (1963) показал, что это не так.

Пример 33.12

В примере 33.10 были найдены квадраты канонических корреляций (33.103). Поскольку в этом примере $n = 3242$, то в силу (33.110) получаем

$$X^2 = 3242\,(0{,}48516 + 0{,}34604 + 0{,}18803) = 3304,$$

что может быть проверено непосредственным вычислением для примера 33.8.

33.52 Существует много (на самом деле бесконечно много) других способов разбиения X^2. Формальная структура таких разбиений дана в **30.44**. Ответ на вопрос о том, интересно ли данное разбиение со статистической точки зрения, зависит от целей анализа. Прежде всего, следует заметить, что величина X^2 сама является компонентой другой величины, которую мы обозначим X^2_T.

Мы отказываемся теперь от предположения об упорядоченности таблиц, однако ограничимся случаем, когда переменные независимы, т. е. $p_{ij} = p_{i\cdot}p_{\cdot j}$ для всех i, j. Вероятности наблюденных частот n_{ij} в этом случае равны

$$\mathsf{P}\{n_{ij} \mid p_{ij}, n\} = \frac{n!}{\prod_{i,j} n_{ij}!} \prod_{i,j} (p_{ij})^{n_{ij}} = \frac{n!}{\prod_{i,j} n_{ij}!} \prod_i p_{i\cdot}^{n_{i\cdot}} \prod_{i,j} \left(\frac{p_{ij}}{p_{i\cdot}}\right)^{n_{\cdot j}} = \tag{33.111}$$

$$= \frac{n! \prod_i p_{i\cdot}^{n_{i\cdot}}}{\prod_i n_{i\cdot}!} \cdot \frac{n! \prod_j p_{\cdot j}^{n_{\cdot j}}}{\prod_j n_{\cdot j}!} \cdot \frac{\prod_i n_{i\cdot}! \prod_j n_{\cdot j}!}{\prod_{i,j} n_{ij}! n!} \tag{33.112}$$

(ср. **33.23** для таблиц 2×2). Левая часть (33.112) и каждый из сомножителей в правой части могут быть приближены распределением χ^2. Записывая

$$n_{ij} = np_{ij} + e_{ij}$$

и используя разложение Стирлинга, находим, что левая часть (33.112) асимптотически представляется в виде

$$\log \mathsf{P}\{n_{ij} \mid p_{ij}, n\} = \text{const} - \tfrac{1}{2} \sum_{i,j} \{e_{ij}^2/(np_{ij})\}.$$

Таким образом, величина

$$X_T^2 = \sum_{i,j} \frac{(n_{ij} - np_{i\cdot}p_{\cdot j})^2}{np_{i\cdot}p_{\cdot j}} \tag{33.113}$$

асимптотически распределена как сумма квадратов rc нормированных нормальных вариант, связанных одним ограничением $\sum_i \sum_j n_{\cdot j} = n$, т. е. как χ^2 с $rc - 1$ степенями свободы.

Аналогично, величина

$$X_R^2 = \sum_i \frac{(n_{i.} - np_{i.})^2}{np_{i.}} \tag{33.114}$$

асимптотически распределена как χ^2_{r-1}, величина

$$X_C^2 = \sum_j \frac{(n_{.j} - np_{.j})^2}{np_{.j}} \tag{33.115}$$

— как χ^2_{c-1}, и, наконец, величина, которая нам уже знакома как «обычный» X^2 и которую мы теперь можем записать в виде

$$X_{RC}^2 = \sum_i \sum_j \frac{(n_{ij} - n_{i.}n_{.j}/n)^2}{n_{i.}n_{.j}/n}, \tag{33.116}$$

асимптотически распределена как χ^2 с $(r-1)(c-1)$ степенями свободы. Соотношения (33.113)—(33.116) дают асимптотическое разбиение

$$X_T^2 = X_R^2 + X_C^2 + X_{RC}^2, \tag{33.117}$$

которое в действительности лишь отражает представление (33.112). Сумма чисел степеней свободы в правой части (33.117) равна числу степеней свободы в левой, т. е.

$$rc - 1 \equiv (r-1) + (c-1) + (r-1)(c-1). \tag{33.118}$$

Таким образом, мы видим (как мы это уже видели в **33.29**), что $(r-1) + (c-1)$ степеней свободы в правой части (33.118) потеряны в «обычном» X^2 (X_{RC}^2 в настоящих обозначениях) вследствие того, что вероятности оцениваются из самой таблицы. Если бы эти вероятности были известны *a priori*, то можно было бы использовать (33.113). Сказанное служит лишь одним из примеров упомянутой в **19.9** потери степеней свободы вследствие необходимости оценивания параметров.

Пример 33.13 (Лэнкастер, 1949b)

Для каждой комбинации признаков A (три категории) и B (три категории) был осуществлен выборочный эксперимент, состоявший фактически в наблюдении пуассоновской случайной величины. Были получены следующие частоты:

3009	2832	3008	8849
3047	3051	2997	9095
2974	3038	3018	9030
9030	8921	9023	26974

Это один из сравнительно нечастых случаев, когда известны априорные маргинальные вероятности. В данном случае $p_{i.} = p_{.j} = 1/3$, $p_{ij} = 1/9$ $(i, j = 1, 2, 3)$. Пользуясь соотношением (33.117), находим:

Компонента	Значение	Степени свободы	Критическое значение при $\alpha = 0{,}05$
X_R^2	3,615	2	5,991
X_C^2	0,828	2	5,991
X_T^2	11,864	8	15,507
$X_T^2 - X_R^2 - X_C^2$	7,421	4	9,488

Ни одно из трех значений X_R^2, X_C^2, X_T^2 не превосходит своей выборочной границы, приведенной в последней колонке. Условия эксперимента, по-видимому, были примерно постоянными.

Если бы мы не знали маргинальных вероятностей, то их пришлось бы оценивать. Мы получили бы $X_{RC}^2 = 7{,}547$ с 4 степенями свободы. Это значение отличается от значения 7,421 для $X_T^2 - X_R^2 - X_C^2$ в приведенной выше таблице, однако разница незначительна. Она появляется вследствие того, что разбиение (33.117) справедливо лишь асимптотически.

33.53 Лэнкастер (1949b), по существу с помощью метода пункта **33.52**, разбивает величину X^2 для таблицы $r \times c$ на $(r-1)(c-1)$ компонент, каждая из которых имеет одну степень свободы. Каждая степень свободы соответствует некоторой классификации 2×2 из этой таблицы. Мы не приводим деталей, однако метод легко понять из следующих двух примеров. Рассмотрим таблицу 2×3

n_{11}	n_{12}	n_{13}	$n_{1.}$
n_{21}	n_{22}	n_{23}	$n_{2.}$
$n_{.1}$	$n_{.2}$	$n_{.3}$	n

Этой таблице, для которой X^2 имеет 2 степени свободы, соответствуют две таблицы 2×2:

n_{11}	n_{12}	$n_{11} + n_{12}$
n_{21}	n_{22}	$n_{21} + n_{22}$
$n_{.1}$	$n_{.2}$	$n_{.1} + n_{.2}$

$(n_{11} + n_{12})$	n_{13}	$n_{1.}$
$(n_{21} + n_{22})$	n_{23}	$n_{2.}$
$(n_{.1} + n_{.2})$	$n_{.3}$	n

(33.119)

Если обычным путем вычислить X^2 для каждой из этих двух таблиц 2×2, то сумма будет приближенно равна значению X^2 для исходной таблицы 2×3.

Аналогично, для таблицы 3×3 с 4 степенями свободы получаем 4 таблицы 2×2:

$$
\begin{array}{cc|c}
n_{11} & n_{12} & n_{11}+n_{12} \\
n_{21} & n_{22} & n_{21}+n_{22} \\
\hline
(n_{11}+n_{21}) & (n_{12}+n_{22}) & (n_{11}+n_{12}+n_{21}+n_{22})
\end{array}
$$

$$
\begin{array}{cc|c}
(n_{11}+n_{12}) & n_{13} & n_{1.} \\
(n_{21}+n_{22}) & n_{23} & n_{2.} \\
\hline
(n_{11}+n_{12}+n_{21}+n_{22}) & (n_{13}+n_{23}) & n_{1.}+n_{2.}
\end{array}
$$

$$
\begin{array}{cc|c}
(n_{11}+n_{21}) & (n_{12}+n_{22}) & (n_{11}+n_{12}+n_{21}+n_{22}) \\
n_{31} & n_{32} & (n_{31}+n_{32}) \\
\hline
n_{.1} & n_{2.} & n_{.1}+n_{.2}
\end{array}
$$

$$
\begin{array}{cc|c}
(n_{11}+n_{12}+n_{21}+n_{22}) & (n_{13}+n_{23}) & n_{.1}+n_{2.} \\
(n_{31}+n_{32}) & n_{33} & n_{3.} \\
\hline
(n_{.1}+n_{.2}) & n_{.3} & n
\end{array}
\qquad (33.120)
$$

Данная процедура является совершенно общей, однако следует помнить, что разбиения не однозначны (поскольку порядок строк и столбцов может вообще быть изменен). Компоненты аддитивны лишь асимптотически, как и в примере 33.13.

Лэнкастер (1949b, 1950) и Ирвин (1949) дают *точный* метод разбиения X^2 на $(r-1)(c-1)$ компонент, соответствующих таблицам 2×2, однако приближенное разбиение уже достаточно хорошо для практических целей; см. упражнение 33.15. Кимбелл (1954) упрощает вычисления, необходимые для точного разбиения.

33.54 Другие типы разбиения X^2 рассматриваются Кокрэном (1954) в обзоре, посвященном таблицам $r \times c$ (а также в действительности и критериям согласия), в котором обсуждаются, кроме того, вопросы возможности анализа таблиц с малыми гипотетическими (при гипотезе независимости) частотами в одной или нескольких клетках таблицы без нарушения χ^2-аппроксимации для X^2. Он дает следующие рекомендации. Если только 1 клетка из 5 или более или 2 клетки из 10 или более имеют гипотетические частоты, меньше 5, то минимальная допустимая гипотетическая частота равна 1. Если таких клеток больше, то обычно

допустима минимальная гипотетическая частота 2, если число степеней свободы меньше 30. Для числа степеней свободы, большего 30, следует пользоваться точными средним и дисперсией, вычисленными Холдейном (1939), и считать X^2 распределенным асимптотически нормально с этими моментами.

Для упорядоченных таблиц Кэнуй (1948) в неопубликованной работе дает разбиение X^2, которое выделяет компоненты линейные, квадратичные и т. д.

Таблицы $2 \times c$: биномиальный критерий однородности

33.55 Частным случаем таблиц $r \times c$, представляющим особый интерес, является таблица $2 \times c$, в которой мы сравниваем c выборок в отношении наличия или отсутствия некоторого свойства. Общая формула (33.62) принимает в этом случае вид

$$X^2 = \sum_{i=1}^{2} \sum_{j=1}^{c} \frac{(n_{ij} - n_{i.}n_{.j}/n)^2}{n_{i.}n_{.j}/n}, \tag{33.121}$$

где $n_{.j}$ — объем j-й выборки и $n = \sum_j n_{.j}$, как и раньше. Полезные точные и приближенные методы вычисления (33.121) содержатся в упражнении 33.21. Если мы обозначим $\hat{p} = n_{1.}/n$ МП-оценку вероятности наблюдения «успеха», вычисленную по данным таблицы, то (33.121) можно будет записать в виде

$$X^2 = \sum_{j=1}^{c} \left[\frac{(n_{1j} - n_{.j}\hat{p})^2}{n_{.j}\hat{p}} + \frac{\{(n_{.j} - n_{1j}) - n_{.j}(1-\hat{p})\}^2}{n_{.j}(1-\hat{p})} \right] = \sum_{j=1}^{c} \frac{(n_{1j} - n_{.j}\hat{p})^2}{n_{.j}\hat{p}(1-\hat{p})}. \tag{33.122}$$

Величина X^2 распределена асимптотически как χ^2 с $c-1$ степенями свободы. Таким образом, критерий однородности c биномиальных выборок, основанный на (33.122), по существу основан на сумме квадратов c биномиальных величин, каждая из которых измеряется от своего математического ожидания (оцененного в предположении независимости) и делится на свою оцененную стандартную ошибку $\{n_{.j}\hat{p}(1-\hat{p})\}^{1/2}$. Число степеней свободы равно $c-1$, так как мы пользуемся оценкой математического ожидания, линейно зависящей от наблюдений. Если бы независимо от наблюдений было известно его значение p, то мы заменили бы $\hat{p}$ на p в (33.122) и получили бы полное число степеней свободы c для X^2.

Армитедж (1955) излагает теорию критериев для обнаружения тренда в исходных вероятностях таблицы $2 \times c$, которая,

по существу, является приложением к этой простой ситуации методов ранжирования и введения меток, рассмотренных в общем случае ранее в этой главе.

Пуассоновский критерий однородности

33.56 Посмотрим теперь, что происходит со статистикой (33.122), когда гипотетическое биномиальное распределение стремится к пуассоновскому в классическом смысле (см. **5.8**), так что $n_{\cdot j} \to \infty$, $p \to 0$, $n_{\cdot j} p \to \lambda$.

В этом случае мы имеем c независимых наблюдений пуассоновской случайной величины n_{1j}. Статистика (33.122) с p вместо $\hat{p}$ принимает вид

$$X^2 = \sum_{j=1}^{c} \frac{(n_{1j} - \lambda)^2}{\lambda} \qquad (33.123)$$

и распределена асимптотически как χ^2 с c степенями свободы (ср. **33.55**). Если, как обычно, надо оценить λ, то пользуются полной достаточной несмещенной оценкой $n_{1\cdot}/c = \bar{n}_1$, являющейся средним c наблюдений. Таким образом, критерий однородности для c пуассоновских величин, основанный на статистике

$$X^2 = \sum_{j=1}^{c} \frac{(n_{1j} - \bar{n}_1)^2}{\bar{n}_1}, \qquad (33.124)$$

имеет $c - 1$ степеней свободы. Одна степень свободы, как и в случае статистики (33.122), теряется в результате оценивания.

Критерии этого и предыдущего пунктов принадлежат Р. А. Фишеру и называются иногда *дисперсионными критериями* для биномиального и пуассоновского распределений. Основанием для этого служит то, что каждый из них основан на сумме c слагаемых, каждое из которых есть отношение квадрата варианты, центрированной относительно ее оцененного среднего, к ее дисперсии, — в случае (33.124) пуассоновские генеральные среднее и дисперсия равны, так что $\bar{n}_1$ служит оценкой для обоих.

33.57 Кокрэн (1954) дает детальное изложение и библиографию для биномиального и пуассоновского дисперсионных критериев и особенно для разбиения степеней свободы X^2 в этих случаях.

Оказывается, в частности, что дисперсионная статистика (33.124) часто приводит к более мощному критерию для проверки гипотезы о том, что выборка извлечена из пуассоновской

совокупности, чем критерий согласия X^2, основанный на группировке наблюдений в соответствии с частотами появления значений 0, 1, 2, ... Основной причиной этого служит то, что для пуассоновского распределения с малыми значениями параметра λ наблюденные частоты ведут себя нерегулярно, начиная с некоторого значения, равного примерно 4 или 5, если λ равно 1 или меньше (см. таблицу 5.3 и пример 19.11). Таким образом, если n не слишком велико, критерий согласия может иметь лишь небольшое число степеней свободы (порядка 5), поскольку большие значения должны быть объединены в один класс, так как критерий можно применять только при достаточно больших гипотетических частотах (см. **30.30**). Эти рассуждения не относятся к дисперсионному критерию, число степеней свободы которого равно $c - 1$ (на единицу меньше числа наблюдений) и который не требует никакой группировки, — этот момент, возможно, не проявился отчетливо при построении критерия через посредство таблицы $2 \times c$, однако он непосредственно следует из (33.124). Таким образом, для «разумных» объемов выборок дисперсионный критерий будет, по-видимому, более мощным.

Поттхофф и Уиттингхилл (1966a, b) приводят другие статистики для проверки однородности биномиальных, мультиномиальных и пуассоновских распределений.

Армитедж (1966) рассматривает моменты дисперсионного критерия X^2 в случае, когда выборка взята из расслоенной совокупности.

Таблицы с несколькими входами

33.58 До сих пор мы всюду, за исключением **33.9—10**, рассматривали связи между двумя категоризованными переменными; наша таблица $r \times c$ была таблицей с двумя входами. Естественно обобщить задачу на случай $p \geqslant 3$ категоризованных переменных. Соответствующая таблица называется *таблицей с несколькими входами* или, иногда, *сложной таблицей сопряженности*. Такие таблицы впервые были рассмотрены К. Пирсоном (1904, 1916) для случая многомерного нормального распределения. Если p-й переменной соответствует r_p категорий, то мы получаем таблицу $r_1 \times r_2 \times \ldots \times r_p$, которая физически может быть представлена только в p-мерном пространстве. В простейшем случае $p = 3$ мы можем представлять таблицу $r_1 \times r_2 \times r_3$ в виде прямоугольного параллелепипеда со строками, столбцами и «слоями», пересечения которых дают клетки таблицы. Чтобы не пользоваться индексами, мы вместо $r_1 \times r_2 \times r_3$ будем писать $r \times c \times l$. В действительности из многомерных таблиц только трехмерной было уделено более чем формальное внимание в литературе, поскольку в случае $p > 3$

не возникает новых теоретических трудностей; однако, как мы увидим, переход от двух измерений к трем требует новых рассмотрений.

Льюис (1962) посвящает данному предмету обширный обзор.

33.59 Рассмотрим сначала подход пункта **33.52**, в котором проводилось разбиение таблицы с двумя входами $r \times c$ при гипотезе независимости. Если обозначить n_{ijk} наблюденную частоту и p_{ijk} вероятность в клетке на пересечении i-й строки, j-го столбца и k-го слоя, то гипотеза о полной независимости запишется в виде

$$H_0\colon\ p_{ijk} = p_{i..}\,p_{.j.}\,p_{..k}, \tag{33.125}$$

где точка, как и прежде, означает суммирование по соответствующему индексу. В случае двух входов мы получили разбиение (33.117), (33.118) на следующие компоненты: «строки», «столбцы» и «строки $\times$ столбцы» с числом степеней свободы $r-1$, $c-1$ и $(r-1)\times(c-1)$ соответственно. В рассматриваемом случае трех входов мы получаем следующие асимптотически аддитивные компоненты:

	Компонента	Число степеней свободы
Строки	X^2_R	$r-1$
Столбцы	X^2_C	$c-1$
Слои	X^2_L	$l-1$
Строки×столбцы	X^2_{RC}	$(r-1)(c-1)$
Строки×слои	X^2_{RL}	$(r-1)(l-1)$
Столбцы×слои	X^2_{CL}	$(c-1)(l-1)$
Строки×столбцы×слои	X^2_{RCL}	$(r-1)(c-1)(l-1)$
Сумма	X^2_T	$rcl-1$

(33.126)

Для таблицы $2\times2\times2$ каждая компонента в (33.126) имеет 1 степень свободы.

Если представлять таблицу $r\times c\times l$ в виде параллелепипеда, то можно сказать, что рассеяние выражается прежде всего в терминах ребер, затем в терминах граней и, наконец, в терминах содержимого таблицы.

Каждая из компонент в (33.126) легко вычисляется. Если имеются гипотетические вероятности для каких-либо (или всех) из $p_{i..}$, $p_{.j.}$, $p_{..k}$, то соответствующие компоненты X^2_R, X^2_C, X^2_L, равны значениям статистики согласия X^2 для маргинальных

распределений строки, столбца и слоя соответственно. Если для какого-то случая неизвестны теоретические вероятности, то соответствующая компонента тождественно равна нулю. Вычислим теперь «обычную» статистику X^2 для проверки независимости в каждой из трех таблиц с двумя входами. Из X^2 для таблицы $(r \times c)$ вычтем $(X_R^2 + X_C^2)$, из X^2 для таблицы $(r \times l)$ вычтем $(X_R^2 + X_L^2)$ и из X^2 для таблицы $(c \times l)$ вычтем $(X_C^2 + X_L^2)$. В результате мы получим X_{RC}^2, X_{RL}^2 и X_{CL}^2 соответственно. Вычислим, наконец, статистику X^2 для проверки независимости в таблице $(r \times c \times l)$, т. е. X_T^2. Статистику X_{RCL}^2 теперь можно получить с помощью вычитания.

Пример 33.14 (Лэнкастер (1951), данные Робертса и др.)

Данные для крыс, классифицированных в соответствии с наличием или отсутствием признаков A, B, D были представлены в виде таблицы $2 \times 2 \times 2$. Как и раньше, мы используем символы α, β, δ для обозначения отсутствия признака.

Имеются следующие частоты:

$$(ABD) = 475, \quad (\alpha BD) = 467,$$
$$(AB\delta) = 460, \quad (\alpha B\delta) = 440,$$
$$(A\beta D) = 462, \quad (\alpha\beta D) = 494,$$
$$(A\beta\delta) = 509, \quad (\alpha\beta\delta) = 427.$$

Мы образуем из них следующие три таблицы 2×2:

Признаки	α	*A*	Сумма
β	921	971	1892
B	907	935	1842
Сумма	1828	1906	3734

Признаки	δ	*D*	Сумма
β	936	956	1892
B	900	942	1842
Сумма	1836	1898	3734

Признаки	δ	*D*	Сумма
α	867	961	1828
A	969	937	1906
Сумма	1836	1898	3734

Гипотетические вероятности всех признаков равны 1/2. Таким образом, для A получаем

$$X_A^2 = \frac{(1828 - 3734/2)^2}{1867} + \frac{(1906 - 3734/2)^2}{1867} = 1{,}6294. \quad (33.127)$$

Подобным же образом мы вычисляем остальные компоненты. Они приведены в третьей колонке следующей таблицы:

Компонента	Число степеней свободы	X^2 с априорными вероятностями	X^2 при оценке параметров
A	1	1,6294	0
B	1	0,6695	0
D	1	1,0295	0
AB	1	0,1296	0,1176
AD	1	4,2517	4,3426
BD	1	0,1296	0,1397
ABD	1	2,7863	2,6904
Сумма	7	10,6256	7,2904

(33.128)

Для одной степени свободы 95-процентная точка распределения χ^2 равна 3,84, а 97,5-процентная точка равна 5,02. Единственной компонентой, близкой к этим значениям, является компонента AD, которая лежит между ними. Следовательно, если вообще существует какая-то связь между факторами, то следует искать связь между A и D. Однако гипотеза о независимости вряд ли нарушена. В самом деле, мы находим $V_{AD} = (X_{AD}^2/n)^{1/2} = 0{,}03$. Во всяком случае компонента ABD лежит внутри выборочных границ, и если бы A и D были связаны, то следовало бы ожидать большего значения компоненты ABD. Кроме того, мы всегда при разбиении X^2 должны помнить, что разделение одного критерия на несколько критериев (здесь на 7) увеличивает вероятность, что какая-нибудь компонента выйдет за свои выборочные границы. В целом, по-видимому, следует считать все три фактора независимыми (или настолько слабо зависимыми, что не имеется определенного указания на взаимозависимость).

Если бы априорные вероятности были не известны, а оценивались по маргинальным частотам, мы получили бы значения X^2, приведенные в последней колонке (33.128). Эти значения, как и должно быть, очень близки к предыдущим, и мы приходим к тому же заключению.

Можно отметить, что мы могли бы иметь априорную информацию о некоторых вероятностях и не иметь ее о других.

В таком случае нам следовало бы оценить неизвестные вероятности и поступать далее, как прежде.

33.60 Природа таблиц со многими входами позволяет рассмотреть большое количество гипотез, отличных от гипотезы о полной независимости, сформулированной в (33.125). С. Рой и Митра (1956), проводя в отношении структуры таблиц с несколькими входами такое же различие, какое мы проводили в **33.18** и **33.31** в отношении таблиц 2×2 и $r \times c$, строят асимптотические критерии X^2 для некоторых таких гипотез. Например, нас может интересовать гипотеза

$$H_1\colon \frac{p_{ijk}}{p_{..k}} = \frac{p_{i.k}}{p_{..k}} \frac{p_{.jk}}{p_{..k}}, \tag{33.129}$$

утверждающая, что в каждом из слоев таблицы с тремя входами переменные, соответствующие строкам и столбцам, независимы. Это аналог равенства нулю частной корреляции между строками и столбцами при фиксированном слое. Или нас может интересовать гипотеза

$$H_2\colon p_{ijk} = p_{ij.} p_{..k} \tag{33.130}$$

утверждающая, что классификация по строкам и столбцам, рассматриваемая как двумерное распределение, не зависит от слоя. Это аналог равенства нулю множественной корреляции между слоями, с одной стороны, и строками и столбцами — с другой.

Проводя суммирование в обеих частях (33.130) сначала по j, затем по i, мы получаем

$$p_{i.k} = p_{i..} p_{..k} \tag{33.131}$$

и

$$p_{.jk} = p_{.j.} p_{..k}. \tag{33.132}$$

Однако только из (33.131), (33.132) не следует (33.130). С. Рой и Кастенбаум (1956) исследовали вопрос о том, какая нужна дополнительная гипотеза, чтобы из (33.131), (33.132) следовало (33.130). Из-за математической трудности они не стали рассматривать естественную гипотезу

$$H_3\colon p_{ijk} = \frac{p_{ij.}}{p_{.j.}} \frac{p_{.jk}}{p_{..k}} \frac{p_{i.k}}{p_{i..}}, \tag{33.133}$$

а вместо нее предложили

$$H_4\colon p_{ijk} = a_{ij} a_{jk} a_{ik}, \tag{33.134}$$

где a — произвольные положительные числа. Они показали, что из (33.134) и (33.131), (33.132) следует (33.130). В упражнении 33.31 дается другое выражение для (33.133).

33.61 В соответствии с терминологией дисперсионного анализа (том 3) (33.134) является гипотезой о равенстве нулю

взаимодействия второго порядка в таблице. Эта задача была впервые рассмотрена для таблицы $2 \times 2 \times 2$ Бартлеттом (1935b), а для таблицы $2 \times 2 \times l$ Нортом (1945). Лэнкастер (1951) предложил иной метод, основанный на компоненте X^2_{RCL} из (33.126). Интерпретация этого метода как критерия для взаимодействия второго порядка критически обсуждается Плэкеттом (1962); см. также полезное обсуждение аналогии с дисперсионным анализом Дэррочем (1962). Плэкетт предложил другой критерий, упрощенный затем Гудмэном (1963b). Гудмэн (1964c) обобщил этот метод на взаимодействия любого порядка, рассмотрел (1963a) случай $2 \times 2 \times l$, дал (1964a) другие методы, основанные на перекрестных отношениях, и дал (1964d) простые методы проверки гипотез и получения доверительных интервалов для взаимодействий второго порядка.

Линдли (1964) проводит байесовский анализ таблиц сопряженности.

33.62 Бёрч (1963) рассматривает МП-оценивание параметров таблицы со многими входами. Он также (1964c) рассматривает критерий существования частной связи в таблице $2 \times 2 \times l$, принадлежащий, по существу, Кокрэну (1954) и основанный на статистике $n_{11\cdot}$, приближенно нормальной со средним и дисперсией, полученными суммированием в (33.51) по l слоям. Если $\theta_k = \log \dfrac{p_{11k}p_{22k}}{p_{12k}p_{21k}}$, то H_0 состоит в том, что $\theta_k = 0$, и критерий будет РНМН против альтернативы H_1: все θ_k равны и положительны. Рассматривается также вопрос об оценивании предполагаемого общего значения θ_k. Проверка гипотезы о *равенстве* θ_k эквивалентна проверке гипотезы о взаимодействии второго порядка (см. **33.61**). Бёрч обобщает теорию на случай таблиц $r \times c \times l$.

Лэнкастер (1960) распространяет на таблицы со многими входами идеи канонического анализа.

УПРАЖНЕНИЯ

33.1 Показать, что коэффициент связи Q больше по абсолютной величине коэффициента коллигации Y, за исключением случая, когда оба равны нулю или единице по абсолютной величине.

33.2 Вывести формулу (33.17) для стандартной ошибки V.

33.3 Для таблицы 3×4 (*Ammon, Zur Anthropologie der Badener*)

Цвет глаз	Цвет волос				Сумма
	B_1	B_2	B_3	B_4	
A_1	1768	807	189	47	2811
A_2	946	1387	746	53	3132
A_3	115	438	288	16	857
Сумма	2829	2632	1223	116	$6800 = n$

показать, что $X^2 = 1075{,}2$ и что, следовательно, коэффициенты (33.63)—(33.65) имеют значения

$$P = 0{,}3695, \qquad T = 0{,}2541, \qquad C = 0{,}2812.$$

33.4 Показать, что для таблицы $r \times c$ коэффициент сопряженности Пирсона P равен коэффициенту Чупрова T для двух значений величины X^2/n, одно из которых равно нулю. Показать, что если X^2/n находится между этими двумя значениями, то $P > T$, а если X^2/n превосходит большее из значений, то $T > P$.

33.5 В следующей таблице приведены результаты экспериментов по иммунизации крупного рогатого скота против туберкулеза.

Таблица 33.4

Данные из отчета об экспериментах Спалинджера (Spahlinger) в Северной Ирландии (1931—1934) (H. M. Stationery Office, 1935)

	Число погибших или серьезно пострадавших от туберкулеза	Число легко пострадавших или непострадавших	Сумма
Привитые вакциной	6	13	19
Непривитые или привитые контрольными средствами	8	3	11
Сумма	14	16	30

Показать, что для этой таблицы при гипотезе о независимости между вакцинацией и восприимчивостью к туберкулезу $X^2 = 4{,}75$, так что гипотеза отвергается при $\alpha \geqslant 0{,}029$; значение α, получающееся при использовании поправки на непрерывность, равно 0,072, а точное значение, полученное методом пунктов **33.19—20**, равно 0, 070.

33.6 Показать, что если объединить две строки и два столбца таблицы $r \times c$, то статистика X^2 для проверки независимости в новой таблице не может быть больше X^2 для прежней таблицы, а вообще будет меньше.

33.7 Показать, что если f — нормированное p-мерное нормальное распределение с матрицей рассеяния $\boldsymbol{V}$ и маргинальными распределениями $f_1, f_2, \ldots, f_p$, то

$$\varphi^2 = \frac{X^2}{n} = \int \ldots \int \frac{(f - f_1 f_2 \ldots f_p)^2}{f_1 f_2 \ldots f_p}\, dx_1 \ldots dx_p = \frac{1}{(|V||W|)^{1/2}} - 1,$$

где $\boldsymbol{W} = 2\boldsymbol{I} - \boldsymbol{V}$.

(К. Пирсон, 1904.)

33.8 Для таблицы с несколькими входами, основанной на классификации нормированного p-мерного нормального распределения с корреляциями ρ_{ij}, показать, что

$$\log(1 + \varphi^2) = -\frac{1}{2}\log|\boldsymbol{I} + \boldsymbol{P}| - \frac{1}{2}\log|\boldsymbol{I} - \boldsymbol{P}|,$$

где величина φ^2 определена в упражнении 33.7, а $\boldsymbol{P}$ — матрица с элементами ρ_{ij}, $i \neq j$, и 0, $i = j$. Пользуясь разложением логарифма, показать, что

$$\varphi^2 > \frac{1}{2} \operatorname{tr} \boldsymbol{P}^2 = \sum_{i<j} \rho_{ij}^2.$$

(Лэнкастер, 1957.)

33.9 Для таблицы $r \times c$ с фиксированными маргинальными частотами рассмотрим статистику

$$\frac{H}{n^2} = \sum_{i=1}^{r} \sum_{j=1}^{c} \left\{ \left(n_{ij} - \frac{n_{i.} n_{.j}}{n} \right)^2 / n_{.j} \right\} = \sum_{i=1}^{r} \frac{n_{i.}}{n} \sum_{j=1}^{c} \frac{D_{ij}^2}{n_{i.} n_{.j}/n},$$

являющуюся взвешенной суммой вкладов строк таблицы в X^2 в (33.62) с весами, равными относительным частотам $n_{i.}/n$. Показать, что при гипотезе независимости

$$\mathbf{M}H = (c-1)n\left(n^2 - \sum_i n_{i.}^2\right) \Big/ (n-1),$$

$$\mathbf{D}H = \frac{4\left\{\left(\sum_i n_{i.}^2\right)^2 - n_i \sum_i n_{i.}^3\right\}}{(n-1)(n-2)} \left\{ \left(c^2 - n \sum_j \frac{1}{n_{.j}} \right)(n-2) - 4(c-1)(n-c) \right\} +$$

$$+ \left[\frac{2n^2 \left(n^2 - \sum_i n_{i.}^2\right)\left(\sum_i n_{i.}^2 - n\right)}{(n-1)^2 (n-2)(n-3)} + \frac{4(n+6)\left\{\left(\sum_i n_{i.}^2\right)^2 - n \sum_i n_{i.}^3\right\}}{(n-1)(n-2)(n-3)} \right] \times$$

$$\times \left\{ n(n-1)\left(c - 1 + \frac{1}{n} - \sum_j \frac{1}{n_{.j}} \right) + (c-1)^2 \right\}.$$

Показать, что если маргинальные суммы строк равны, так что $n_{i.} = n/r$, то $H = \dfrac{n^2}{r} X^2$ и

$$\mathbf{M}H = \frac{n^3}{(n-1)r}(r-1)(c-1),$$

$$\mathbf{D}H = \frac{2n^5 (n-r)(r-1)}{(n-1)^2 (n-2)(n-3) r^2} \times$$

$$\times \left\{ n(n-1)\left((c-1) + \frac{1}{n} - \sum_j \frac{1}{n_{.j}} \right) + (c-1)^2 \right\},$$

так что

$$\mathbf{M}(X^2) \to (r-1)(c-1), \qquad \mathbf{D}(X^2) \to 2(r-1)(c-1),$$

как это и должно быть, поскольку предельное распределение X^2 есть χ^2 с $(r-1)(c-1)$ степенями свободы.

(Смит, 1951, 1952; см. также Холдейн, 1939.)

33.10 Подобно тому, как это было сделано в примере 33.8, показать для таблицы 33.5, что $\sum_{i,j} a_{ij}b_{ij} = 2 \times 13\,264\,256$, и получить отсюда значения $t_b = 0{,}658$, $t_c = 0{,}633$.

Таблица 33.5

7477 женщин 30—39 лет, служащих Королевских артиллерийских заводов Великобритании (1943—1946): дальность зрения невооруженным глазом

Правый глаз / Левый глаз	Высшая степень	Вторая степень	Третья степень	Низшая степень	Сумма
Высшая степень	1520	266	124	66	1976
Вторая степень	234	1512	432	78	2256
Третья степень	117	362	1772	205	2456
Низшая степень	36	82	179	492	789
Сумма	1907	2222	2507	841	7477

33.11 Для таблиц 4×4 примера 33.8 и упражнения 33.10 показать, что коэффициент G, определенный формулой (33.79), принимает соответственно значения 0,776 и 0,798 с максимальными стандартными ошибками (33.80), равными 0,022 и 0,014. Показать также, что значения t_c, равные 0,629 и 0,633, имеют максимальные стандартные ошибки 0,029 и 0,019 соответственно.

(Гудмэн и Крускал, 1963.)

33.12 Следующие данные (Чепмэн) связывают условия, при которых выполнялось домашнее задание (имеющие градации от лучших A_1 до худших A_5), и оценку качества работы учителем (от лучшей B_1 до худшей B_3):

	A_1	A_2	A_3	A_4	A_5	Сумма
B_1	141	67	114	79	39	440
B_2	131	66	143	72	35	447
B_3	36	14	38	28	16	132
Сумма	308	147	295	179	90	1019

Показать, сопоставляя категориям в качестве меток натуральные числа, что коэффициент регрессии качества домашнего задания в зависимости от условий его выполнения (0,025 в этих единицах) находится внутри обычных границ для выборочных флуктуаций относительно нуля (стандартная ошибка равна 0,016).

(Иэйтс, 1948.)

33.13 Приведенная ниже таблица (A. R. Treloar) связывает состояние зубов 135 женщин со средним ежедневным потреблением кальция каждой из них:

		Среднее потребление кальция в день (в граммах)			
		0–0,40	0,40–0,55	0,55–0,70	свыше 70
Состояние зубов	*A*	5	3	10	11
	B	4	5	8	6
	C	26	11	3	6
	D	23	11	1	2

Показать, что канонические корреляции равны

$$R_1 = 0{,}56273, \qquad R_2 = 0{,}10869, \qquad R_3 = 0{,}00045,$$

так что

$$nR_1^2 = 42{,}74984,$$
$$nR_2^2 = 1{,}59497,$$
$$nR_3^2 = 0{,}00003$$
$$X^2 = 44{,}34484.$$

Только первая из этих компонент X^2 превосходит общепринятые критические значения для χ^2. Показать далее, что канонические метки, соответствующие R_1, равны:

Состояние зубов	Потребление кальция
A: −1,3880	0–0,40: 0,8397
B: −1,0571	0,40–0,55: 0,4819
C: 0,6016	0,55–0,70: −1,5779
D: 0,9971	свыше 0,70: −1,1378

Отметим, в частности, изменение в поведении меток, соответствующих потреблению кальция, в точке 0,70, которое подтверждает вызванное данными впечатление о том, что существует верхний предел, после которого увеличение потребления кальция не улучшает состояния зубов.

(Э. Уилльямс, 1952.)

33.14 В примере 33.10 показать, что

$$\boldsymbol{N'N} = \begin{Vmatrix} 0{,}633768533 & 0{,}192522708 & 0{,}146101456 & 0{,}092877193 \\ 0{,}192522708 & 0{,}444704410 & 0{,}241885812 & 0{,}093129969 \\ 0{,}146101456 & 0{,}241885812 & 0{,}474670557 & 0{,}200085907 \\ 0{,}092877193 & 0{,}093129969 & 0{,}200085907 & 0{,}466094141 \end{Vmatrix}$$

и что из четырех характеристических корней этой матрицы один равен единице, а остальные имеют значения (33.103). Использовать (33.98) для непосредственного вычисления меток y (33.105), а затем найти метки x (33.104), используя (33.99).

33.15 Для данных примера 33.13 показать, что четырем таблицам 2×2 (33.120) соответствуют компоненты

$$\begin{matrix} 2{,}860 & 2{,}180 \\ 2{,}526 & 0{,}005, \end{matrix}$$

сумма которых равна 7,571, тогда как для исходной таблицы 3×3 $X^2 = 7{,}547$.

(Лэнкастер, 1949b).

33.16 Использовать в **33.53** метод **33.52** для доказательства асимптотического разбиения X^2 для таблицы 2×3 на компоненты (33.119) и показать, что доказательство может быть обобщено на случай таблицы $r \times c$.

(Лэнкастер, 1949b.)

33.17 Для случая мультиномиального распределения

$$(p_1 + p_2 + \ \dots \ + p_r)^n$$

определим

$$w_i = \frac{n_i - np_i}{\{np_i\,(1 - p_i)\}^{1/2}}, \qquad i = 1, 2, \dots, r.$$

Показать, что в обозначениях частных корреляций и регрессий

$$\rho_{ij} = -\{p_i p_j/(1 - p_i)\,(1 - p_j)\}^{1/2},$$

$$\rho_{ij.kl\,\dots\,m} = -(p_i p_j/[\{(1 - p_i) - (p_k + p_l + \ \dots \ + p_m)\}\{(1 - p_j) - (p_k + p_l + \ \dots \ + p_m)\}])^{1/2},$$

$$\sigma_i^2 = 1, \qquad \sigma_{i.j}^2 = \{1 - (p_i + p_j)\}/\{(1 - p_i)\,(1 - p_j)\},$$

$$\beta_{ij.kl\,\dots\,m} = -\{p_i p_j\,(1 - p_j)/(1 - p_i)\}^{1/2}/\{1 - p_j\,(p_k + p_l + \ \dots \ + p_m)\}.$$

Показать также, что w_1/σ_1, $w_{2\cdot1}/\sigma_{2\cdot1}$, $w_{3\cdot21}/\sigma_{3\cdot21}$ и т. д. асимптотически нормально распределены с нулевыми средними, единичными дисперсиями и нулевыми корреляциями. Показать далее, что $w_{k\cdot12\,\dots\,(k-1)} = 0$ и что, следовательно, первые $k - 1$ величин w дают разбиение X^2 на $k - 1$ независимых компонент.

(Лэнкастер, 1949b.)

33.18 Пусть $\boldsymbol{C}$ и $\boldsymbol{D}$ — ортогональные матрицы с рангами r и c соответственно, и пусть $(rc \times rc)$-матрица $\boldsymbol{K}$ есть их прямое произведение $\boldsymbol{C} \times \boldsymbol{D}$. Пусть переменные

$$x_{ij} = (n_{ij} - np_{i.}p_{.j})/(np_{i.}p_{.j})^{1/2}$$

образуют вектор-столбец

$$\boldsymbol{x} \equiv \{x_{11}, x_{12}, \dots, x_{1c}, x_{21}, x_{22}, \dots, x_{2c}, \dots, x_{r1}, x_{r2}, \dots, x_{rc}\}'.$$

Показать, что при подходящем выборе элементов $\boldsymbol{C}$ и $\boldsymbol{D}$ матрица $\boldsymbol{Y} = \boldsymbol{Kx}$ дает компоненты X^2 для таблицы сопряженности $r \times c$, причем y_{11} соответствует общей сумме, y_{1k} $(k \neq 1)$ — суммам по столбцам, y_{k1} $(k \neq 1)$ — суммам по строкам, а оставшиеся элементы — остальным $(r - 1)\,(c - 1)$ степеням свободы (см. **33.52**).

(Лэнкастер, 1951.)

33.19 В частном случае таблицы 3×3 в упражнении 33.18 положить

$$\boldsymbol{C} = \boldsymbol{D} = \begin{Vmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \end{Vmatrix}.$$

Для данных примера 33.13 найти матрицу значений y

$$\begin{pmatrix} 0 & -0{,}812829 & -0{,}409012 \\ -1{,}834459 & 1{,}653094 & 1{,}471167 \\ -0{,}499424 & -1{,}587173 & -0{,}070020 \end{pmatrix}.$$

33.20 Показать, что для таблицы $2 \times c$ с $n_{1\cdot} = n_{2\cdot}$ статистика X^2 принимает вид

$$X^2 = \sum_{j=1}^{c} \frac{(n_{1j} - n_{2j})^2}{n_{1j} + n_{2j}}.$$

Это статистика критерия для проверки однородности двух выборок равного объема, классифицированных в соответствии с c категориями.

33.21 В таблице $2 \times c$ предположим, что $n_{1\cdot} > n_{2\cdot}$, и выберем небольшие целые числа k, h, $k > h \geqslant 1$, такие, что k/h служит приближением для $n_{1\cdot}/n_{2\cdot}$ и, следовательно, разность $(hn_{1\cdot} - kn_{2\cdot})$ мала. Показать, что для этой таблицы (33.121) можно записать точно следующим образом:

$$X^2 = \frac{n^2}{(h+k)^2 n_{1\cdot} n_{2\cdot}} \left\{ \sum_{j=1}^{c} (hn_{1j} - kn_{2j})^2 / n_{\cdot j} - (hn_{1\cdot} - kn_{2\cdot})^2 / n \right\},$$

а приближенно, с относительной ошибкой, не превосходящей $1 - \dfrac{hn_{1\cdot} - kn_{2\cdot}}{n}$, в виде

$$X^2 \approx \frac{1}{hk} \left\{ \sum_{j=1}^{c} (hn_{1j} - kn_{2j})^2 / n_{\cdot j} - (hn_{1\cdot} - kn_{2\cdot})^2 / n \right\}.$$

(Холдейн, 1955b.)

33.22 Для таблицы 2×11

25	80	38	52	9	21	33	24	30	51	56	419
1	15	12	8	0	7	6	2	7	7	3	68
26	95	50	60	9	28	39	26	37	58	59	487

использовать результаты последнего упражнения с $h = 1$, $k = 6$. Показать, что точная формула дает для X^2 значение 16,709, а приближенная — значение 16,393. Ошибка этого приближения равна 1,9%, тогда как верхняя граница этой ошибки равна 2,3%.

(Холдейн, 1955b.)

33.23 Показать (см. упражнение 33.20), что в таблице сопряженности $r \times r$ можно проверить гипотезу о полной симметрии (т. е. H_0: $p_{ij} = p_{ji}$, $i, j = 1, 2, \ldots, r$, где p_{ij} — вероятность появления наблюдения в i-й строке и j-м столбце) с помощью статистики

$$u = \sum_{i<j}\sum \frac{(n_{ij} - n_{ji})^2}{n_{ij} + n_{ji}},$$

асимптотически распределенной как χ^2 с $\frac{1}{2} r(r-1)$ степенями свободы.

(Боукер, 1948.)

33.24 Показать, что в таблице $r \times r$ с однородными категоризациями (см. 33.2) можно проверять однородность двух маргинальных распределений (т. е. H_0: $p_{i\cdot} = p_{\cdot i}$, $i = 1, 2, \ldots, r$) с помощью квадратичной формы от $r-1$ асимптотически нормальных переменных $d_i = (n_{i\cdot} - n_{\cdot i})$, $i = 1, 2, \ldots, r-1$,

$$Q = \boldsymbol{d}'\boldsymbol{V}^{-1}\boldsymbol{d} = \sum_{i=1}^{r-1}\sum_{j=1}^{r-1} V^{ij} d_i d_j,$$

где V^{ij} есть (i, j)-й элемент матрицы, обратной $(r-1)\times(r-1)$-матрице рассеяния $\boldsymbol{V}$ величин d_i, элементы которой имеют вид

$$V_{ii} = n_{i\cdot} + n_{\cdot i} - 2n_{ii}, \qquad V_{ij} = -(n_{ij} + n_{ji}), \qquad i \neq j.$$

Показать, что Q асимптотически имеет распределение χ^2 с $r-1$ степенями свободы.

(Стьюарт, 1955b; Мадански (1963) рассматривает критерий отношения правдоподобия для маргинальной однородности в случае таблицы с несколькими входами.)

33.25 Показать, что в таблице $r \times r$ гипотеза

$$H_0\colon \frac{p_{ii}}{p_{i\cdot}p_{\cdot i}} = \frac{p_{jj}}{p_{j\cdot}p_{\cdot j}}, \qquad i, j = 1, 2, \ldots, r,$$

может быть проверена с помощью статистики

$$V = \sum_{i=1}^{r} n_{i\cdot}n_{\cdot i} \frac{\left\{\dfrac{n_{ii}}{n_{i\cdot}n_{\cdot i}} - \dfrac{\sum_i n_{ii}}{\sum_i n_{i\cdot}n_{\cdot i}}\right\}^2}{\sum_i n_{ii} \Big/ \sum_i n_{i\cdot}n_{\cdot i}},$$

распределенной асимптотически как χ^2 с $r-1$ степенями свободы.

(Дёрбин; см. Гласс (1954), стр. 234.)

33.26 Проверить значения компонент X^2 в двух разложениях в (33.128), используя метод, приведенный после (33.126).

33.27 Перегруппируем частоты примера 33.13, кроме 3018, в таблицу $2 \times 2 \times 2$

3009	2832	3008	2974
3047	3051	2997	3038

(цифры справа соответствуют «слою», лежащему выше, а слева — лежащему ниже). Считая гипотетические вероятности равными 1/2, получить следующее разбиение X^2:

	X^2 (гипотетические вероятности)	X^2 (параметры, оцененные по данным)
R	4,0115	0
C	1,1503	0
L	0,2540	0
RC	2,7357	2,7824
RL	1,7372	1,7547
LC	1,3524	1,3607
RCL	0,4690	0,5107
	11,7101	6,4085

(Лэнкастер, 1951.)

33.28 Пусть в таблице $2 \times 2 \times 2$

$$T = (p_{1..})^{1/2}, \quad U = (p_{.1.})^{1/2}, \quad V = (p_{..1})^{1/2},$$

$$t = (p_{2..})^{1/2}, \quad u = (p_{.2.})^{1/2}, \quad v = (p_{..2})^{1/2},$$

и пусть матрица $\boldsymbol{M}$ порядка (8×8) является прямым произведением матриц

$$\begin{pmatrix} V & v \\ -v & V \end{pmatrix} \times \begin{pmatrix} U & u \\ -u & U \end{pmatrix} \times \begin{pmatrix} T & t \\ -t & T \end{pmatrix}.$$

Пусть, кроме того,

$$x_{ijk} = (n_{ijk} - np_{i..}p_{.j.}p_{..k})/(np_{i..}p_{.j.}p_{..k})^{1/2}$$

и $\boldsymbol{x}$ — вектор-столбец

$$\boldsymbol{x} \equiv (x_{111}, x_{211}, x_{121}, x_{221}, x_{112}, x_{212}, x_{122}, x_{222})'.$$

Показать, что элементы $\boldsymbol{Mx}$ служат компонентами (асимптотически независимых) случайных величин χ для (тот же порядок; ср. (33.126))

$$T, R, C, RC, L, RL, CL, RCL.$$

(Лэнкастер, 1951.)

33.29 Маргинальные вероятности $p_{i.}$, $p_{.j}$ таблицы $r \times c$ известны, и требуется оценить клеточные вероятности p_{ij}. Дана выборка из n наблюдений, взятых из мультиномиального распределения с вероятностями p_{ij}. Показать, что МП-оценки $\hat{p}_{ij}$ для p_{ij} находятся как решения системы $(r-1)(c-1)$ уравнений

$$\frac{n_{ij}}{p_{ij}} - \frac{n_{ic}}{p_{ic}} - \frac{n_{rj}}{p_{rj}} + \frac{n_{rc}}{p_{rc}} = 0, \qquad i = 1, 2, \ldots, r-1; \qquad j = 1, 2, \ldots, c-1.$$

Показать также, что эти оценки являются модифицированными несмещенными МД-оценками для p_{ij}, получаемыми применением (19.59) к n_{ij} $(i = 1, 2, \ldots, r-1;\ j = 1, 2, \ldots, c-1)$; при этом в точной матрице рассеяния $\boldsymbol{V}$ p_{ij} всюду заменяются на $\hat{p}_{ij}$.

(Эль-Бадри и Стефан, 1955; см. также Дж. Смит, 1947.)

33.30 В c различных ситуациях одни и те же n индивидов наблюдались с целью выяснения наличия (значение 1) или отсутствия (значение 0) некоторого признака. Результаты были представлены в виде таблицы $2 \times 2 \ldots \ldots \times 2 = 2^c$. Обозначим T_j ($j = 1, 2, \ldots, c$) общее число единиц в j-й ситуации и u_i ($i = 1, 2, \ldots, 2^c$) — число единиц среди c координат i-й клетки таблицы. Показать, что если вероятность «единицы» одинакова во всех c ситуациях, то статистика

$$Q = c(c-1) \sum_{j=1}^{c} (T_j - \bar{T})^2 \Big/ \left(c \sum_i u_i - \sum_i u_i^2 \right)$$

(где суммирование в знаменателе ведется по всем непустым клеткам) асимптотически распределена как χ^2 с $c - 1$ степенями свободы.

(Кокрэн, 1950; Мадански (1963) обобщает этот результат на случай таблиц r^c.)

33.31 Определяя отношение $R_{ij\cdot} = p_{ij\cdot}/(p_{i\cdot\cdot}p_{\cdot j\cdot})$ и, аналогично, $R_{i\cdot k}$, $R_{\cdot jk}$, показать, что гипотеза (33.133) может быть записана в виде

$$H_3\colon\ R_{ij\cdot} = (p_{ijk}/p_{\cdot\cdot k})/\{(p_{i\cdot k}/p_{\cdot\cdot k})(p_{\cdot jk}/p_{\cdot\cdot k})\}$$

с аналогичными выражениями для H_3 в терминах $R_{i\cdot k}$ и $R_{\cdot jk}$, и показать, что H_3 является гипотезой о том, что каждое из этих трех отношений инвариантно относительно изменения опущенной категоризованной переменной.

ГЛАВА 34

ПОСЛЕДОВАТЕЛЬНЫЕ МЕТОДЫ

Последовательные процедуры

34.1 В предыдущих главах, рассматривая задачи, связанные с выборками, мы обычно предполагали объем выборки n фиксированным. Это могло происходить по разным причинам: или мы выбирали n заранее; или вообще не могли выбирать n, например, когда нам представлены результаты законченного эксперимента; или это могло быть обусловлено тем, что объем выборки определялся некоторым другим критерием, например, если было решено проводить наблюдения в течение данного промежутка времени. Мы делали наши выводы, оставаясь в области, где n постоянно. Например, находя стандартную ошибку некоторой оценки, мы делаем вероятностные утверждения, относящиеся к множеству выборок объема n. Можно было бы сказать, что наши формулы являются условными относительно n. Ясно, что если n определяется некоторым способом, не связанным со значениями наблюдений, то такая трактовка справедлива.

34.2 Иногда, однако, объем выборки представляет собой случайную величину, зависящую от значений наблюдений. С одним из таких простейших случаев мы уже сталкивались в примере 9.13 (том 1, стр. 313). Предположим, что мы последовательно выбираем людей, чтобы определить относительное число имеющих редкую группу крови. Вместо выбора, скажем, 1000 индивидуумов и подсчета числа представителей с этой группой крови, мы можем предпочесть продолжать выбор до появления 20 таких случаев. Мы увидим позже, почему эта процедура может быть предпочтительней, а на данном этапе примем на веру, что она заслуживает рассмотрения. Производя выборки по такой схеме, мы, несомненно, нашли бы, что число n, требуемое для достижения фиксированного числа успехов, скажем 20, заметно колеблется. Оно должно быть не меньше 20, но может быть и бесконечным (хотя вероятность бесконечного продолжения выборки равна нулю, так что почти наверное мы должны рано или поздно остановиться).

34.3 Процедуры, подобные этой, называются *последовательными*. Их типичной особенностью является *выборочная схема*, устанавливающая правило, по которому мы решаем на каждом этапе, прекратить или продолжать выбор. В нашем примере правило было очень простым: если испытание дало неблагоприятный исход, то продолжаем выбор; если достигнут успех, то также продолжаем выбор, за исключением случая, когда в предыдущих испытаниях достигнуто 19 успехов: тогда выбор прекращается. Решение в любой момент, вообще говоря, зависит от наблюдений, сделанных до этого момента. Таким образом, для последовательности значений $x_1, x_2, \ldots, x_n$ объем выборки, при котором мы прекращаем выбор, зависит от иксов. Именно этот факт является характерной особенностью последовательного анализа.

Последовательные методы начали развиваться во время второй мировой войны в основном Вальдом (работа которого полностью изложена в его книге, Вальд (1947)) в США и одновременно Барнардом (1946) в Англии.

34.4 Обычная ситуация, когда мы фиксируем объем выборки заранее, может рассматриваться как очень специальный случай последовательной схемы. В этом случае выборочная процедура такова: продолжать выбор до получения n членов независимо от того, какие наблюдаются значения. Однако это настолько вырожденный специальный случай, что в нем фактически теряется суть последовательного метода.

Если процедура останавливается с вероятностью единица, то схему называют *замкнутой*. Если же существует ненулевая вероятность того, что выбор может продолжаться неограниченно, то схема называется *открытой*. В этой главе мы не будем серьезно рассматривать открытые схемы. Они имеют, очевидно, меньшее практическое значение по сравнению с замкнутыми схемами, и их обычно приходится редуцировать к замкнутым схемам, устанавливая верхний предел для объема выборки. Такое усечение часто делает трудным точное определение их свойств.

В литературе по этому вопросу не всегда единообразно используются термины. «Замкнутые» иногда означает «усеченные», так говорят применительно к схемам, в которых определенное правило замыкания устанавливает верхний предел объема выборки. Соответственно «открытые» иногда означает «неусеченные».

Пример 34.1

В качестве примера довольно простой последовательной схемы рассмотрим выбор из (большой) популяции с долей успехов ω. Мы будем продолжать выбор до получения m успехов,

после чего остановимся. Вряд ли нужно доказывать, что такая схема замкнута. Вероятность того, что в бесконечной последовательности не появится m успехов, равна нулю.

Вероятность $m-1$ успехов в первых $n-1$ испытаниях совместно с успехом в m-м испытании (см. **5.14—15**) равна

$$\binom{n-1}{m-1}\omega^m\chi^{n-m}, \quad n=m,\ m+1,\ \ldots, \tag{34.1}$$

где $\chi=1-\omega$. Это есть распределение объема выборки n. Производящая функция частот для n (с началом в нуле) дается выражением

$$\left(\frac{\omega t}{1-\chi t}\right)^m. \tag{34.2}$$

Отсюда получаем производящую функцию семиинвариантов:

$$\psi(t)=\log\left(\frac{\omega e^t}{1-\chi e^t}\right)^m=m\log\left(\frac{\omega}{e^{-t}-\chi}\right).$$

Разлагая это выражение до члена t^2, находим

$$\varkappa_1=\frac{m}{\omega}, \tag{34.3}$$

$$\varkappa_2=\frac{m(1-\omega)}{\omega^2}=\frac{m\chi}{\omega^2}. \tag{34.4}$$

Таким образом, среднее значение объема выборки n равно m/ω. Из этого не следует, что m/n является несмещенной оценкой для ω. В действительности такая несмещенная оценка дается выражением

$$p=(m-1)/(n-1), \tag{34.5}$$

потому что

$$\mathbf{M}\left(\frac{m-1}{n-1}\right)=\sum_{n=m}^{\infty}\left(\frac{m-1}{n-1}\right)\binom{n-1}{m-1}\omega^m\chi^{n-m}=$$

$$=\omega\sum_{m}^{\infty}\binom{n-2}{m-2}\omega^{m-1}\chi^{n-m}=\omega. \tag{34.6}$$

Дисперсия этой оценки в сжатой форме не выражается. В самом деле,

$$\mathbf{M}\left(\frac{m-1}{n-1}\right)^2=(m-1)\,\omega^{m-1}\chi^{1-m}\sum_{n=m}^{\infty}\binom{n-2}{m-2}\frac{\chi^{n-1}}{n-1}=$$

$$=(m-1)\,\omega^m\chi^{1-m}\int_0^{\chi}\sum_{n=m}^{\infty}\binom{n-2}{m-2}t^{n-2}\,dt=$$

$$=(m-1)\,\omega^m\chi^{1-m}\int_0^{\chi}t^{m-2}(1-t)^{1-m}\,dt. \tag{34.7}$$

Полагая $u = \omega t/\{\chi(1-t)\}$, находим

$$\mathsf{M}\left(\frac{m-1}{n-1}\right)^2 = (m-1)\,\omega^2 \int_0^1 \frac{u^{m-2}\,du}{\omega+\chi u} =$$

$$= (m-1)\,\omega^2 \int_0^1 u^{m-2}\left\{\sum_{j=0}^{\infty}\chi^j(1-u)^j\right\}du =$$

$$= (m-1)\,\omega^2 \sum_{j=0}^{\infty}\chi^j \mathrm{B}(m-1,\, j+1) =$$

$$= \omega^2\left[1 + \frac{\chi}{m} + \frac{2!\chi^2}{m(m+1)} + \frac{3!\chi^3}{m(m+1)(m+2)} + \ldots\right]. \qquad (34.8)$$

Отсюда, вычитая ω^2, имеем

$$\mathsf{D}p = \frac{\omega^2\chi}{m}\left[1 + \frac{2\chi}{m+1} + \frac{6\chi^2}{(m+1)(m+2)} + \ldots\right]. \qquad (34.9)$$

Можно получить несмещенную оценку для $\mathsf{D}p$ в простой и замкнутой форме. Тем же путем, каким мы пришли к (34.6), показывается, что

$$\mathsf{M}\,\frac{(m-1)(m-2)}{(n-1)(n-2)} = \omega^2.$$

Следовательно,

$$\mathsf{M}\left\{\left(\frac{m-1}{n-1}\right)^2 - \frac{(m-1)(m-2)}{(n-1)(n-2)}\right\} = \mathsf{M}\left(\frac{m-1}{n-1}\right)^2 - \omega^2.$$

Таким образом,

$$\widehat{\mathsf{D}p} = \left(\frac{m-1}{n-1}\right)^2 - \frac{(m-1)(m-2)}{(n-1)(n-2)} = \frac{(m-1)(n-m)}{(n-1)^2(n-2)} =$$

$$= \frac{p(1-p)}{n-2} = \frac{p^2q}{m-q}. \qquad (34.10)$$

Заметим, что для больших n это выражение асимптотически совпадает с соответствующим результатом для выборок фиксированного объема n.

Оценка коэффициента вариации величины p для данного отрицательно-биномиального распределения имеет вид

$$\frac{p(m-q)^{1/2}}{\{p^2(1-p)\}^{1/2}},$$

и для малых p она приближенно равна $\sqrt{m-1}$. Таким образом, для последовательного процесса *относительная* выборочная вариация величины p приближенно постоянна.

Эта последовательная схема, часто называемая *обратным* биномиальным выбором, впервые обсуждалась Холдейном (1945) и Финни (1949b). Найт (1965) сделал эту теорию единой для биномиального, пуассоновского, гипергеометрического и экспоненциального распределений.

34.5 Выборка по качественному признаку занимает такое большое место в последовательном анализе, что, прежде чем переходить к более общим рассмотрениям, мы обсудим полезный графический метод представления процесса выбора.

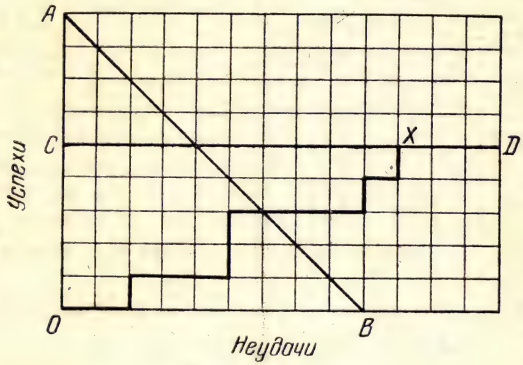


Рис. 34.1.

Возьмем координатную сетку, как на рис. 34.1, и будем откладывать число неудач на оси абсцисс, а число успехов — на оси ординат.

Последовательный выбор может быть представлен на этой сетке траекторией, выходящей из начала координат, причем сдвиг вправо на один шаг соответствует неудаче $Н$, а один шаг вверх отвечает успеху $У$. Например, траектория OX изображает последовательность $ННУННННУУННННУНУ$. На этой диаграмме правило остановки эквивалентно некоторому барьеру. Например, линия AB задается уравнением $У + Н = 9$ и, таким образом, соответствует случаю выборки фиксированного объема $n = 9$. Линия CD соответствует $У = 5$ и, следовательно, определяет правило остановки такого типа, как рассмотренное в примере 34.1, с $m = 5$. Траектория OX, изображающая выборку объема 15, является тогда одной из траекторий, которая заканчивается в X. Если X представляет собой точку с координатами (x, y), то число различных траекторий, соединяющих O с X, равно числу способов, которыми x может быть выбрано из $(x + y)$. Вероятность достижения точки X есть вероятность осуществления x успехов и y неудач, а именно

$$\binom{x+y}{x} \omega^x \chi^y.$$

Пример 34.2. Разорение игрока

Одна из старейших задач теории вероятностей относится к последовательным процессам. Рассмотрим двух игроков A и B и серию игр, в каждой из которых A выигрывает с вероятностью ω и B — с вероятностью $1-\omega$. Проигравший в каждой игре платит победителю рубль. Если A начинает с a рублями, а B с b рублями, то каковы вероятности их разорения (игрок считается разорившимся, когда он проигрывает свой последний рубль)?

Серии последовательных игр, подобные этой, могут быть представлены на графике, аналогичном рис. 34.1. Мы можем считать успехом победу игрока A в одной игре. Игра продолжается до тех пор, пока и A и B имеют деньги, и прекращается, когда A имеет $a+b$. (т. е. B проиграл всю свою начальную сумму) или когда B имеет $a+b$ рублей (A проиграл свою начальную сумму). Поэтому границами, определяющими выборочную схему, являются линии $y-x=-a$ и $y-x=b$.

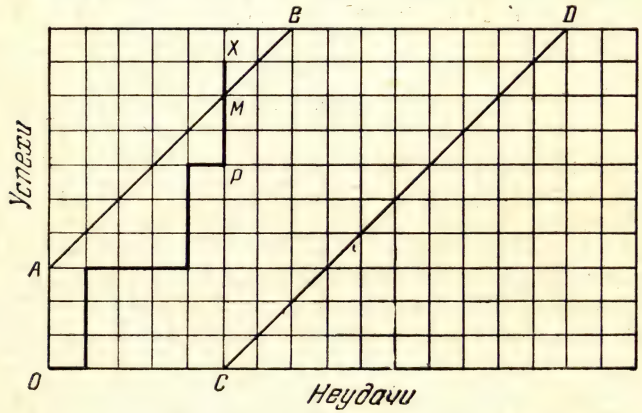


Рис. 34.2.

Рисунок 34.2 служит иллюстрацией для случая $a=5$, $b=3$. Линии AB, CD наклонены под 45° к осям и проходят через точки с координатами $H=0$, $У=3$ и $H=5$, $У=0$ соответственно. Для любой точки между этими линиями $У-H$ меньше, чем 3, и $H-У$ меньше, чем 5. На линии AB $У-H=3$, и если траектория достигает этой линии, то B, проиграв на три игры больше, чем A, разоряется; аналогично, когда траектория достигает CD, то разоряется игрок A. Таким образом, последовательная схема выглядит так: если точка лежит между линиями, то выбор продолжается; если достигнута линия AB, то вы-

бор прекращается и разоряется игрок B; при достижении границы CD следует остановка и разорение игрока A.

Легко получить интересующие нас вероятности. Пусть u_x есть вероятность того, что, имея x рублей, игрок A разорится. Рассматривая следующую игру, выводим уравнение

$$u_x = \omega u_{x+1} + \chi u_{x-1} \tag{34.11}$$

с граничными условиями

$$u_0 = 1, \quad u_{a+b} = 0. \tag{34.12}$$

Обшее решение уравнения (34.11) есть

$$u_x = At_1^x + Bt_2^x,$$

где t_1 и t_2 — корни квадратного уравнения

$$\omega t^2 - t + \chi = 0,$$

а именно:

$$t = 1 \quad \text{и} \quad t = \chi/\omega.$$

Предположим, что $\omega \neq \chi$. Тогда, используя (34.12), находим решение

$$u_x = \frac{(\chi/\omega)^{a+b} - (\chi/\omega)^x}{(\chi/\omega)^{a+b} - 1}, \qquad \omega \neq \frac{1}{2}. \tag{34.13}$$

Если же $\omega = 1/2$, то решение имеет вид

$$u_x = \frac{a+b-x}{a+b}. \tag{34.14}$$

В частности, в начале игры, когда $x = a$, при $\omega = 1/2$

$$u_a = \frac{b}{a+b}. \tag{34.15}$$

34.6 Подобные ситуации можно, очевидно, обобщать многими способами и, в частности, можно устанавливать различного типа границы. При этом замкнутая схема — это, по существу, такая схема, для которой достоверно, что граница будет достигнута.

Пусть, например, схема задается так: если проигрывает A, то он платит один рубль, если же проигрывает B, то он платит k рублей. Траектория на рис. 34.2, соответствующая такой схеме, состоит тогда из шагов единичной длины, параллельных оси абсцисс, и шагов длины k, параллельных оси ординат. Этот пример дает нам возможность выделить момент, который постоянно вносит трудность в математическую сторону последовательных схем: траектории могут не заканчиваться точно на границе, а пересекать ее. Например, для $k = 3$ такой траекто-

рией является ломаная OX на рис. 34.2. После двух успехов и пяти неудач мы приходим в точку P. Следующий успех приведет нас в точку X, попадая в которую мы пересекаем границу в точке M. Конечно, мы останавливаемся на том шаге, при котором граница достигается или пересекается. Существенный момент данного примера заключается в том, что точная вероятность достижения границы в точке M равна нулю. Фактически эта точка недостижима. Позднее мы увидим, что из-за такого отсутствия непрерывности иногда бывает трудно установить точные и краткие утверждения об интересующих нас вероятностях. Мы будем называть такие ситуации «граничными эффектами». В большинстве практически случаев ими можно пренебрегать.

Последовательные критерии для проверки гипотез

34.7 Применим идеи последовательного анализа к задачам проверки гипотез и в первую очередь к различению двух гипотез H_0 и H_1. Предположим, что эти гипотезы относятся к параметру θ, который может принимать значения θ_0 и θ_1 соответственно, т. е. H_0 и H_1 являются простыми гипотезами. Нам нужно найти выборочную схему, делящую выборочное пространство на три взаимно непересекающиеся области: (а) область ω_a, при попадании в которую выборочной точки мы принимаем гипотезу H_0 (отвергаем H_1); (б) область ω_r, при попадании в которую выборочной точки мы принимаем H_1 (и отвергаем H_0); (в) остальная часть выборочного пространства, ω_c, — когда точка попадает в эту область, мы продолжаем выбор.

Если в примере 34.2 принять разорение игрока A за H_0, а разорение игрока B за H_1, то областью ω_a будет область справа от CD, включая саму линию; областью ω_r является область выше AB, включая саму линию; областью ω_c служит область между линиями.

Оперативная характеристика

34.8 Вероятность принятия гипотезы H_0, когда верна H_1, есть функция от θ_1. Обозначим ее $K(\theta_1)$. Если схема замкнута, то вероятность отвергнуть гипотезу H_0, когда H_1 верна, равна $1 - K(\theta_1)$. Как функция от θ_1 для различных значений θ_1, она является просто функцией мощности. Подобно предыдущему, мы могли бы, конечно, все делать в терминах мощности, но в последовательном анализе стало традицией иметь дело с самой функцией $K(\theta_1)$.

$K(\theta)$, рассматриваемая как функция от θ, называется «оперативной характеристикой» (ОХ) схемы. Откладывая ее по оси ординат, а θ — по оси абсцисс, получим «кривую оперативной характеристики», являющуюся дополнением (до единицы) функции мощности.

Средний объем выборки

34.9 Второй функцией, используемой для характеристики последовательного критерия, является «средний объем выборки» (СОВ). Это есть математическое ожидание объема выборки n, требуемого для решения о принятии гипотезы H_0 или H_1 и, следовательно, для прекращения выбора. ОХ для H_0 и H_1 не зависит от объема выборки, а зависит только от постоянных, заданных выборочной схемой в самом начале, до выбора. СОВ измеряет объем наблюдений, требуемый при применении данной схемы.

Пример 34.3

Рассмотрим выборку из (большой) совокупности качественных признаков с долей успехов ω, и пусть ω мало. Нас интересует возможность того, что ω меньше некоторого заданного значения ω_0. Такая задача часто встречается, например, в ситуации, где производитель некоторого изделия желает гарантировать, что доля отказов в партии изделий не превосходит некоторого фиксированного числа. Прежде всего рассмотрим альтернативу $\omega_1 > \omega_0$.

Возьмем очень простую схему. Если успех не появляется, то мы продолжаем выбор до наперед заданного объема выборки n_0 и принимаем значение ω_0. Если же, однако, появился успех, то мы принимаем ω_1 и прекращаем работу.

Пусть ω является истинной вероятностью успеха. Тогда вероятность того, что мы примем гипотезу, равна $(1-\omega)^{n_0} = \chi^{n_0}$. Это есть ОХ. Она представляет собой J-образную кривую, монотонно убывающую от $\omega = 0$ до $\omega = 1$. Два наших конкретных значения даются просто ее ординатами в точках ω_0 и ω_1.

Общий смысл ситуации требует, чтобы мы принимали меньшее из ω_0 и ω_1, если успех не появился, и большее, если успех появился. Пусть ω_0 — меньшее число. Тогда вероятность ошибки первого рода α равна $1-\chi_0^{n_0}$, а вероятность ошибки второго рода β равна $\chi_1^{n_0}$. Если мы поменяем местами ω_0 и ω_1, то α-ошибка будет $1-\chi_1^{n_0}$, а β-ошибка $\chi_0^{n_0}$, и они обе больше, чем в предыдущем случае.

В этом частном случае, используя ОХ, можно получить критерий для проверки сложной гипотезы H_0: $\omega \leqslant \omega_0$ против

H_1: $\omega > \omega_0$. Действительно, если $\omega < \omega_0$, то вероятность α-ошибки меньше, чем $1 - \chi_0^{n_0}$, а вероятность β-ошибки меньше, чем $\chi_0^{n_0}$.

СОВ находится вычислением среднего значения объема выборки m, на котором мы остановимся. Очевидно, что для любого заданного ω оно равно

$$\sum_{m=1}^{n_0-1} m\omega(1-\omega)^{m-1} + n_0(1-\omega)^{n_0-1} =$$

$$= -\omega\frac{\partial}{\partial\omega}\sum_{0}^{n_0-1}(1-\omega)^m + n_0(1-\omega)^{n_0-1} = \frac{1-(1-\omega)^{n_0}}{\omega}. \qquad (34.16)$$

В данном случае СОВ также является убывающей функцией от ω, так как он равен

$$\frac{1-\chi^{n_0}}{1-\chi} = 1 + \chi + \chi^2 + \ldots + \chi^{n_0-1}.$$

Мы видим, что СОВ будет разным в зависимости от того, какое из значений ω_0 или ω_1 является истинным.

Для дискретных распределений трудно произвести сравнение результатов последовательной процедуры с результатами, относящимися к обыкновенной выборке фиксированного объема, особенно это касается сравнения ошибок первого и второго рода. Рассмотрим, однако, случай $\omega_0 = 0{,}1$ и $n = 30$. Из таблиц биномиального распределения (например, *Biometrika Tables*, таблица 37) мы видим, что вероятность 5 или более успехов равна приблизительно 0,18. Таким образом, основываясь на выборке фиксированного объема 30, мы можем отвергнуть значение $\omega = 0{,}1$ с ошибкой первого рода 0,18. При альтернативном значении $\omega = 0{,}2$ вероятность 4 или менее успехов равна 0,26, что и является ошибкой второго рода.

Для последовательного критерия с объемом выборки n_0 ошибка первого рода равна $1 - \chi_0^{n_0}$, а ошибка второго рода равна $\chi_1^{n_0}$. При объеме выборки 2 ошибка первого рода равна 0,19, а ошибка второго рода равна 0,64. При объеме выборки 6 ошибки равны 0,47 и 0,26 соответственно. Ясно, что в этом простом примере мы не можем установить соответствие между двумя типами ошибок, но очевидно, что для достижения фиксированного уровня ошибки любого рода в последовательном случае потребуется меньший объем выборки. Пользуясь более гибкой последовательной схемой, мы можем обеспечить заданные уровни ошибок обоих родов при меньшем СОВ, чем в случае выборки фиксированного объема. На самом деле в малости среднего объема выборки и состоит одно из основных преимуществ последовательных методов (см. пример 34.10).

Критерий отношения вероятностей Вальда

34.10 Предположим, что мы выбираем одно за другим m значений $x_1, x_2, \ldots, x_m$ из совокупности $f(x, \theta)$. На любом этапе отношение вероятностей данной выборки для гипотез H_0 $(\theta = \theta_0)$ и H_1 $(\theta = \theta_1)$ равно

$$L_m = \prod_{i=1}^{m} f(x_i, \theta_1) \Big/ \prod_{i=1}^{m} f(x_i, \theta_0). \qquad (34.17)$$

С помощью способа, который будет описан позднее, выберем два числа A и B, соответствующих желаемым α- и β-ошибкам, и установим следующий последовательный критерий: выбор продолжается, пока $B < L_m < A$; принимаем H_1, когда впервые $L_m \geqslant A$; принимаем H_0, когда впервые $L_m \leqslant B$.

Эквивалентной, но более удобной для вычислений функцией является логарифм L_m. Теперь критическое неравенство имеет вид

$$\log B < \sum_{i=1}^{m} \log f(x_i, \theta_1) - \sum_{i=1}^{m} \log f(x_i, \theta_0) < \log A. \qquad (34.18)$$

Такие критерии мы будем называть «последовательными критериями отношения вероятностей» (ПКОВ).

34.11 Нам часто будет удобно обозначать

$$z_i = \log \{f(x_i, \theta_1)/f(x_i, \theta_0)\}. \qquad (34.19)$$

Тогда критическое неравенство (34.18) будет эквивалентно утверждению относительно накопленных сумм z_i. Докажем прежде всего, что ПКОВ заканчивается с вероятностью единица, т. е. что он замкнут.

Выборка завершается, если

$$\sum z_i \geqslant \log A \quad \text{или} \quad \sum z_i \leqslant \log B.$$

Случайные величины z_i независимы с дисперсиями, скажем, $\sigma^2 > 0$. Тогда $\sum_{i=1}^{m} z_i$ имеет дисперсию $m\sigma^2$. С ростом m дисперсия увеличивается и вероятность того, что значение $\sum z_i$ остается между конечными пределами $\log B$ и $\log A$, стремится к нулю. Более точно, распределение среднего $\bar{z}$ стремится, в силу центральной предельной теоремы, к (нормальному) распределению с дисперсией σ^2/m, и следовательно, вероятность того, что это среднее попадет между $(\log B)/m$ и $(\log A)/m$, стремится к нулю.

Стейн (1946) показал, что $\mathbf{M}(e^{mt})$ существует для любого комплексного числа t, действительная часть которого меньше некоторого $t_0 > 0$. Отсюда следует, что случайная величина m имеет моменты всех порядков

Пример 34.4

Рассмотрим снова биномиальное распределение с вероятностью успеха ω. Если имеется k успехов в первых m испытаниях, то ПКОВ задается статистикой

$$\log L_m = k \log \frac{\omega_1}{\omega_0} + (m-k) \log \frac{1-\omega_1}{1-\omega_0}. \qquad (34.20)$$

Эту величину мы вычисляем, проводя выбор, который продолжается до тех пор, пока не достигнем граничных значений $\log B$ или $\log A$. Теперь покажем, как определяются числа A и B.

34.12 Замечательный факт состоит в том, что числа A и B могут быть получены очень просто (по крайней мере с приемлемой точностью приближения) из вероятностей ошибок первого и второго рода α и β без знания исходного распределения. Таким образом, не требуется решать задачи, связанные с распределением. Это не означает, что последовательный процесс является свободным от распределения. Зная распределение, мы подставляем его в критерий L_m (34.17) и далее обращаемся непосредственно с этим отношением правдоподобия. И неудивительно, что ПКОВ обладают определенными оптимальными свойствами, поскольку они используют всю имеющуюся в распоряжении информацию, включая порядок появления выборочных значений.

Рассмотрим выборку, для которой L_m лежит между A и B при первых $n-1$ испытаниях, а затем становится $\geqslant A$ в n-м испытании, так что мы принимаем H_1 (и отвергаем H_0). По определению вероятность получения такой выборки по крайней мере в A раз больше при гипотезе H_1, чем при гипотезе H_0. Поскольку это верно для любой точки выборки, то это верно и для совокупности всевозможных выборок, ведущих к принятию H_1. Вероятность принятия H_1, когда выполняется H_0, равна α, а вероятность принятия H_0, когда выполняется H_1, равна $1-\beta$. Следовательно, $1-\beta \geqslant A\alpha$, или

$$A \leqslant \frac{1-\beta}{\alpha}. \qquad (34.21)$$

Подобным образом, рассматривая случаи, в которых принимается H_0, мы получим, что $\beta \leqslant B(1-\alpha)$, или

$$B \geqslant \frac{\beta}{1-\alpha}. \qquad (34.22)$$

34.13 Если бы наши границы были такими, что значения A и B достигались бы точно, т. е. если бы не было граничного эффекта, то мы могли бы написать

$$A = \frac{1-\beta}{\alpha}, \quad B = \frac{\beta}{1-\alpha}. \qquad (34.23)$$

В действительности Вальд (1947) показал, что для всех прак-

тических целей можно предполагать эти равенства выполненными. Пусть

$$a = \frac{1-\beta}{\alpha}, \qquad b = \frac{\beta}{1-\alpha}, \tag{34.24}$$

и пусть истинные ошибки первого и второго рода для границ a и b равны α' и β'. Тогда из (34.21) имеем

$$\frac{\alpha'}{1-\beta'} \leqslant \frac{1}{a} = \frac{\alpha}{1-\beta}, \tag{34.25}$$

а из (34.22)

$$\frac{\beta'}{1-\alpha'} \leqslant b = \frac{\beta}{1-\alpha}. \tag{34.26}$$

Следовательно,

$$\alpha' \leqslant \frac{\alpha(1-\beta')}{1-\beta} \leqslant \frac{\alpha}{1-\beta}, \tag{34.27}$$

$$\beta' \leqslant \frac{\beta(1-\alpha')}{1-\alpha} \leqslant \frac{\beta}{1-\alpha}. \tag{34.28}$$

Кроме того,

$$\alpha'(1-\beta) + \beta'(1-\alpha) \leqslant \alpha(1-\beta') + \beta(1-\alpha'),$$

или

$$\alpha' + \beta' \leqslant \alpha + \beta. \tag{34.29}$$

На практике α и β малы, чаще всего они равны 0,01 или 0,05. Из (34.27) и (34.28) следует, что величины, на которые α' может превосходить α или β' может превосходить β, ничтожны. Более того, из (34.29) мы видим, что или $\alpha' \leqslant \alpha$, или $\beta' \leqslant \beta$. Следовательно, худшее, что может быть при использовании a и b вместо A и B, — это увеличение одной из ошибок, и то лишь на очень малую величину. Такая процедура поэтому будет всегда безопасна в том смысле, что для всех практических целей не происходит увеличения ошибок неправильного решения. Избегая утомительных повторений, мы будем впредь использовать равенства (34.23), если только не оговорено противное.

Пример 34.5

Рассмотрим снова биномиальный закон из примера 34.4 с $\alpha = 0{,}01$, $\beta = 0{,}10$, $\omega_0 = 0{,}01$ и $\omega_1 = 0{,}03$. Имеем для k успехов и $n - k$ неудач (логарифмируя по основанию 10)

$$\log \frac{\beta}{1-\alpha} \leqslant (n-k) \log \frac{1-\omega_1}{1-\omega_0} + k \log \frac{\omega_1}{\omega_0} \leqslant \log \frac{1-\beta}{\alpha},$$

или

$$\log \frac{10}{99} \leqslant (n-k) \log \frac{97}{99} + k \log 3 \leqslant \log 90,$$

или $-0{,}995653 \leqslant -0{,}0088635\,(n-k) + (0{,}477121)\,k \leqslant 1{,}954243$. Деля на 0,0088635, находим с точностью до ближайших целых

$$-112 \leqslant 54k - n \leqslant 220. \tag{34.30}$$

Пользуясь таким критерием, мы, например, должны принять ω_0, если не появилось неудачного исхода в первых 112 испытаниях. Если одна неудача произошла в 100-м испытании, а другая в 200-м, то мы не можем принять решение прежде 220-го (т. е. $112 + (2 \times 54)$) испытания. И если в 200 испытаниях произошло 6 неудач, скажем в 50-м, 100-м, 125-м, 150-м, 175-м, 200-м, то мы также не можем отклонить гипотезу, так как $54k - n$ равно 124 в 200 испытаниях; но если эксперимент был затем продолжен и величина $54k - n$ превзошла 220, то мы должны принять ω_1.

OX последовательного критерия отношения вероятностей

34.14 Рассмотрим функцию

$$L^h = \left\{ \frac{f(x, \theta_1)}{f(x, \theta_0)} \right\}^h, \tag{34.31}$$

где h зависит от θ. Функция $g(x, \theta) = L^h f(x, \theta)$ есть функция плотности при любом значении θ, если только

$$\mathbf{M}(L^h) = \int \left\{ \frac{f(x, \theta_1)}{f(x, \theta_0)} \right\}^h f(x, \theta)\, dx = 1. \tag{34.32}$$

Можно показать (см. упражнение 34.4), что существует не более одного ненулевого значения h, удовлетворяющего этому уравнению. Рассмотрим следующее правило: принимаем H_0, продолжаем выбор или принимаем H_1 согласно неравенству

$$B^h \leqslant \frac{\prod \{L^h f(x, \theta)\}}{\prod \{f(x, \theta)\}} \leqslant A^h. \tag{34.33}$$

Оно, очевидно, эквивалентно обычному правилу (34.18) при условии, что $h > 0$. Рассмотрим проверку гипотезы H: истинное распределение есть $f(x, \theta)$ — против альтернативы G: истинное распределение есть $g(x, \theta)$. Если α' и β' — ошибки, а отношение правдоподобия таково, как в (34.33), то

$$A^h = \frac{1 - \beta'}{\alpha'}, \qquad B^h = \frac{\beta'}{1 - \alpha'}, \tag{34.34}$$

и, следовательно,

$$\alpha' = \frac{1 - B^h}{A^h - B^h}.$$

Так как α' есть функция мощности, когда выполняется H_1, то ее дополнение, OX, дается равенством

$$1 - \alpha' = K(\theta) = \frac{A^h - 1}{A^h - B^h}. \tag{34.35}$$

Такая же формула выполняется, если $h < 0$.

Теперь мы можем найти OX критерия. Когда $h(\theta) = 1$, $\theta = \theta_0$, а когда $h(\theta) = -1$, $\theta = \theta_1$. Для других значений мы должны решить (34.32) относительно θ и найденное значение затем подставить в (34.35). Но на самом деле, чтобы построить кривую оперативной характеристики $K(\theta)$, в этом нет необходимости. Мы можем взять само $h(\theta)$ в качестве параметра и строить зависимость (34.35) относительно него.

Пример 34.6

Рассмотрим еще раз биномиальный закон из предыдущего примера. Для дискретных значений 1 (успех) и 0 (неудача) мы можем написать

$$f(1, \omega) = \omega, \qquad f(0, \omega) = 1 - \omega.$$

Тогда (34.32) принимает вид

$$\omega\left(\frac{\omega_1}{\omega_0}\right)^h + (1 - \omega)\left(\frac{1 - \omega_1}{1 - \omega_0}\right)^h = 1,$$

или

$$\omega = \frac{1 - \left(\frac{1 - \omega_0}{1 - \omega_1}\right)^h}{\left(\frac{\omega_1}{\omega_0}\right)^h - \left(\frac{1 - \omega_1}{1 - \omega_0}\right)^h}. \tag{34.36}$$

Из (34.35) при $A = (1 - \beta)/\alpha$ и $B = \beta/(1 - \alpha)$ имеем

$$K(\omega) = \frac{\left(\frac{1 - \beta}{\alpha}\right)^h - 1}{\left(\frac{1 - \beta}{\alpha}\right)^h - \left(\frac{\beta}{1 - \alpha}\right)^h}. \tag{34.37}$$

Мы можем теперь начертить $K(\omega)$ относительно ω, используя (34.36) и (34.37) как параметрические уравнения от h.

СОВ для последовательного критерия отношения правдоподобия

34.15 Рассмотрим последовательность из n случайных величин z_i. Если бы n было фиксировано, то мы получили бы

$$\mathsf{M}\left(\sum_{i=1}^{n} z_i\right) = n\mathsf{M}(x).$$

Для последовательного выбора это неверно, но взамен имеет место результат

$$\mathbf{M}\left(\sum_{i=1}^{n} z_i\right) = \mathbf{M}(n)\,\mathbf{M}(z), \tag{34.38}$$

который совсем не так очевиден, как кажется. Этот результат принадлежит Вальду и Блекуэллу (1946), а приводимое ниже доказательство предложено Джонсоном (1959b).

Пусть каждое z_i имеет среднее μ, $\mathbf{M}|z_i| \leqslant C < \infty$, и пусть вероятность того, что n примет значение k, равна P_k $(k = 1, 2, \ldots)$. Введем «маркирующую» переменную y_i, которая равна единице, если z_i наблюдалось (т. е. если $n \geqslant i$), и нулю в противном случае. Тогда

$$\mathbf{P}(y_i = 1) = \mathbf{P}(n \geqslant i) = \sum_{j=i}^{\infty} P_j. \tag{34.39}$$

Пусть теперь $Z_n = \sum_{i=1}^{n} z_i$. Имеем

$$Z_n = \sum_{i=1}^{\infty} y_i z_i,$$

$$\mathbf{M}(Z_n) = \mathbf{M} \sum_{i=1}^{\infty} (y_i z_i) = \sum_{i=1}^{\infty} \mathbf{M}(y_i z_i), \tag{34.40}$$

и $\mathbf{M}(Z_n)$ конечно, если $\mathbf{M}(n)$ конечно. Кроме того, так как y_i зависит только от $z_1, z_2, \ldots, z_{i-1}$ и не зависит от z_i, то

$$\mathbf{M}(y_i z_i) = \mathbf{M}(y_i)\,\mathbf{M}(z_i). \tag{34.41}$$

Следовательно,

$$\mathbf{M}(Z_n) = \sum \mathbf{M}(y_i)\,\mathbf{M}(z_i) = \mu \sum \mathbf{M}(y_i) =$$

$$= \mu \sum_{i=1}^{\infty} \{P_i + P_{i+1} + \ldots\} = \mu \sum_{i=1}^{\infty} i P_i = \mu \mathbf{M}(n),$$

откуда вытекает (34.38).

Из последнего равенства находим

$$\mathbf{M}(n) = \frac{\mathbf{M}(Z_n)}{\mathbf{M}(z)}. \tag{34.42}$$

С принятой нами степенью аппроксимации выбор прекращается, когда Z_n принимает одно из двух значений: $\log A$ с вероятностью $1 - K(\theta)$ и $\log B$ с вероятностью $K(\theta)$. Таким образом,

$$\mathbf{M}(n) = \frac{K \log B + (1 - K) \log A}{\mathbf{M}(z)}, \tag{34.43}$$

что является приближенной формулой для среднего объема выборки.

Пример 34.7

В биномиальном случае получаем

$$\mathbf{M}(z) = \mathbf{M} \log\left(\frac{\omega_1}{\omega_0}\right) = \omega \log \frac{\omega_1}{\omega_0} + (1-\omega) \log \frac{1-\omega_1}{1-\omega_0}. \quad (34.44)$$

СОВ может быть тогда вычислен по формуле (34.43), если заданы ω_1, ω_2, A и B (или α и β). Он, конечно, зависит от ω.

34.16 Для практически применений последовательная проверка качественных признаков часто представляется так, чтобы вычисления производились в целых числах. Неравенство (34.30) относится к рассматриваемому случаю. Мы можем переписать его в виде

$$332 \geqslant 220 + (n-k) - 53k \geqslant 0.$$

Можно представить себе игру, в которой начальная сумма равна 220. Если произошла неудача, мы прибавляем единицу к этой сумме; если произошел успех, мы вычитаем 53 единицы. Игра прекращается, как только сумма падает до нуля или возрастает до 332, что отвечает принятию значений ω_0 или ω_1 соответственно.

34.17 В такой схеме предположим, что начальная сумма равна S_2. При каждой неудаче мы выигрываем единицу, а при каждом успехе теряем b единиц. Если сумма возрастает на S_1, достигая значения $S_1 + S_2$ (равного, скажем, $2S$), мы принимаем одну гипотезу; если она упадет до нуля, мы примем другую гипотезу. Пусть сумма в какой-либо точке равна x, а u_x — вероятность того, что сумма достигнет значения $2S$, не обращаясь до этого в нуль. Рассмотрим результат следующего испытания. Неудача увеличивает сумму до $x+1$, а успех уменьшает ее до $x-b$. Таким образом, имеем уравнение

$$u_x = (1-\omega)\, u_{x+1} + \omega u_{x-b} \quad (34.45)$$

с начальными условиями

$$u_0 = u_{-1} = u_{-2} = \ldots = u_{-b+1} = 0, \quad (34.46)$$

$$u_{2S} = 1. \quad (34.47)$$

При $b = 1$ это уравнение легко решить, как в примере 34.2. При $b > 1$ (мы предполагаем его целым) решение более громоздко. Мы приведем без доказательства решение, полученное Бёрмэном (1946):

$$u_x = \frac{F(x)}{F(2S)}, \quad (34.48)$$

где

$$\left.\begin{aligned} F(x) &= \chi^{-x}\left\{1 - \binom{x-b-1}{1}\omega\chi^b + \right. \\ &+ \left.\binom{x-2b-1}{2}(\omega\chi^b)^2 + \binom{x-3b-1}{3}(\omega\chi^b)^3 + \ldots\right\}, \quad x > 0, \\ F(x) &= 0, \qquad x \leqslant 0. \end{aligned}\right\} \tag{34.49}$$

Здесь ряд продолжается до тех пор, пока $x - kb - 1 > 0$. Бёрмэн нашел также выражения для СОВ и для дисперсии объема выборки.

34.18 Энскомб (1949а) табулировал функции этого вида. Положив

$$R_1 = \frac{S_1}{b+1}, \qquad R_2 = \frac{S_2}{b+1}, \tag{34.50}$$

Энскомб табулировал R_1, R_2 при некоторых значениях ошибок α, β (на самом деле $1-\alpha$ и β) и отношения S_1/S_2. Им приведены также значения $\omega(b+1)$.

Если заданы ω_0, ω_1, α, β, то мы можем найти R_1 и R_2 в зависимости от значения отношения S_1/S_2. Так, для $\omega_0 = 0{,}01$, $\omega_1 = 0{,}03$, $S_2 = 2S_1$, $\alpha = 0{,}01$, $\beta = 0{,}10$ находим приближенно $R_2 = 4$, $R_1 = 2$. Кроме того, $\omega(b+1) = 0{,}571$, или $b = 56$. Затем из (34.50) получаем $S_1 = 114$, $S_2 = 228$. Согласие с примером 34.5 ($S_1 = 112$, $S_2 = 220$, $b = 53$) вполне удовлетворительное. СОВ для $\omega = 0{,}01$ равен 253, а для $\omega = 0{,}03$ он равен 306.

34.19 Поучительно рассмотреть, что происходит в пределе, когда единицы 1 и b становятся малыми по сравнению с общей суммой $2S$. На рис. 34.1 мы можем представить это как сжимание сетки таким образом, чтобы ступенчатые линии приближались к непрерывным случайным траекториям частиц, подвергающихся бесконечно малым возмущениям в двух взаимно перпендикулярных направлениях. С этой точки зрения данная задача связана с теорией броуновского движения и диффузии. Если Δ — разностный оператор, определенный соотношением

$$\Delta u_x = u_{x+1} - u_x,$$

то мы можем написать уравнение (34.45) в виде

$$\{(1-\omega)(1+\Delta) + \omega(1+\Delta)^{-b} - 1\}u_x = 0. \tag{34.51}$$

Для малых b последнее почти эквивалентно уравнению

$$\left\{(1-\omega) + \omega - 1 + (1-\omega)\Delta - b\omega\Delta + \frac{1}{2}b(b+1)\omega\Delta^2\right\}u_x = 0,$$

или

$$\left\{(1-\omega-b\omega)\Delta+\frac{1}{2}b(b+1)\omega\Delta^2\right\}u_x=0. \tag{34.52}$$

В пределе получаем

$$\frac{d^2u}{dx^2}+\lambda\frac{du}{dx}=0, \tag{34.53}$$

где

$$\lambda=\frac{2(1-\omega-b\omega)}{\omega b(b+1)}. \tag{34.54}$$

Общим решением уравнения (34.53) является

$$u_x=k_1+k_2e^{-\lambda x},$$

и, учитывая граничные условия $u_{2S}=1$, $u_0=0$, имеем

$$u_x=\frac{1-\exp(-\lambda x)}{1-\exp\{-\lambda(S_1+S_2)\}}. \tag{34.55}$$

Таким образом, для $x=S_2$ вероятность принятия равна

$$u_{S_2}=\frac{\exp(\lambda S_2)-1}{\exp(\lambda S_2)-\exp(-\lambda S_1)}. \tag{34.56}$$

34.20 Как и прежде, напишем $R_1(b+1)=S_1$, $R_2(b+1)=S_2$, и пусть ω стремится к нулю так, чтобы $\gamma=\omega(b+1)$ оставалось ограниченным. Из (34.54) видим, что λ стремится к нулю, но

$$b\lambda\to\delta=\frac{2\{1-\omega(b+1)\}}{\omega(b+1)}=\frac{2(1-\gamma)}{\gamma}. \tag{34.57}$$

Тогда λS_1 стремится к $S_1\delta/b$, т. е. к $R_1\delta$, если только $\gamma\neq 1$, и (34.56) превращается в

$$u_{S_2}=\frac{\exp(R_2\delta)-1}{\exp(R_2\delta)-\exp(-R_1\delta)}, \qquad \gamma\neq 1. \tag{34.58}$$

При δ, стремящемся к нулю,

$$u_{S_2}=\frac{R_2}{R_2+R_1}, \qquad \gamma=1. \tag{34.59}$$

(34.58) можно сопоставить с формулой (34.37), которая для малых δ может быть записана в виде

$$u_{S_2}=\frac{\exp\left(\frac{1-\beta}{\alpha}h\right)-1}{\exp\left(\frac{1-\beta}{\alpha}h\right)-\exp\left(-\frac{1-\alpha}{\beta}h\right)}. \tag{34.60}$$

34.21 Использование последовательных методов в контроле качества изделий привело к существенному развитию результатов такого типа, как в **34.17—18**. У нас здесь нет возможности детально обсуждать указанные результаты, а читатель, который этим интересуется, должен обратиться к какому-нибудь руководству по контролю качества. Мы же просто приведем некоторые из обобщений предшествующей теории в качестве иллюстрации ее области применения.

(а) *Правила остановки.* Даже для замкнутых схем может быть желательным остановить выборку на некотором этапе. Например, обстоятельства могут препятствовать выбору после некоторого момента времени или в клинических испытаниях медицинская профессиональная этика может потребовать замены лекарства новым, выглядящим перспективно, прежде, чем его достоинства будут полностью установлены. Усечение последовательных схем может производиться различными способами, простейшие из которых состоят в остановке либо при достижении заданного объема выборки, либо по истечении данного времени. В таких случаях наши общие понятия оперативной характеристики и среднего объема выборки остаются без изменений, но математические и вычислительные детали обычно значительно усложняются. Армитедж (1957) рассматривал последовательный выбор при различных ограничениях.

(б) *Обследование с заменой всех дефектных изделий годными.* В схемах, которые мы рассматривали, гипотезы были таковы, что исследуемая партия или совокупность принималась или отвергалась при наличии определенной доли некоторого качественного признака. Если этот признак есть «дефектность», то мы можем предпочесть не отвергать партию в целом, а проверить каждый ее член и заменить дефектные. Это само по себе не влияет на общий характер схемы (решение отвергнуть заменяется решением исправить), но это оказывает, конечно, влияние на долю отказов во всей совокупности партий, к которым выборочный план применяется, и следовательно, воздействует на значение параметров, входящих в план. Указанная доля отказов во всей совокупности партий известна под названием среднего выходного уровня качества. Данная теория была изучена Бартки (1943).

(в) *Двойная выборка.* Развивая дальше эту идею, можно обнаружить, что ее выгоднее осуществлять в несколько этапов. Например, мы можем иметь четыре возможных решения: принять; отвергнуть; продолжать выборку; отложить суждение, но полностью проверить. Очевидно, здесь имеется большое разнообразие возможных действий. Превосходным примером является процедура двойной выборки Доджа и Ромига (1944). С этой идеей мы снова встретимся в **34.36**.

Пример 34.8

Для нормального распределения с единичной дисперсией рассмотрим проверку гипотезы о среднем H_0: $\mu = \mu_0$ против H_1: $\mu = \mu_1$. Определяя z, как в (34.19), имеем

$$\left.\begin{aligned} z_i &= -\tfrac{1}{2}(x_i-\mu_1)^2+\tfrac{1}{2}(x_i-\mu_0)^2=(\mu_1-\mu_0)\,x_i-\tfrac{1}{2}(\mu_1^2-\mu_0^2),\\ Z_m &= \sum_{i=1}^{m} z_i = m(\mu_1-\mu_0)\,\bar{x}-\tfrac{1}{2}m(\mu_1^2-\mu_0^2). \end{aligned}\right\} \tag{34.61}$$

Мы принимаем H_0 или H_1, когда эта величина $\leqslant \log B$ или $\geqslant \log A$ соответственно. Для оперативной характеристики получаем из (34.35)

$$K(\mu) = (A^h-1)/(A^h-B^h), \tag{34.62}$$

где h находится из уравнения

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left[h\left\{(\mu_1-\mu_0)\,x-\tfrac{1}{2}(\mu_1^2-\mu_2^2)\right\}\right]\exp\left\{-\tfrac{1}{2}(x-\mu)^2\right\}dx=1, \tag{34.63}$$

которое, как легко видеть, эквивалентно уравнению

$$\exp\{\mu^2-h\mu_1^2+h\mu_0^2-(\mu-h\mu_1+h\mu_0)^2\}=1,$$

из которого следует

$$h=(\mu_1+\mu_0-2\mu)/(\mu_1-\mu_0), \quad \mu_1\neq\mu_0, \quad h\neq 0. \tag{34.64}$$

Вычисляя теперь h из (34.64) для различных значений μ и подставляя его в (34.62), мы можем начертить кривую оперативной характеристики.

Аналогично, обращаясь к СОВ, имеем

$$\mathbf{M}(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left\{-\tfrac{1}{2}(x-\mu)^2\right\}\left\{(\mu_1-\mu_0)\,x-\tfrac{1}{2}(\mu_1^2-\mu_0^2)\right\}dx=$$

$$=(\mu_1-\mu_0)\,\mu-\tfrac{1}{2}(\mu_1^2-\mu_0^2). \tag{34.65}$$

Вновь для различных значений μ можно определить СОВ из равенств (34.62), (34.65) и

$$\mathbf{M}(n)=\frac{K\log B+(1-K)\log A}{\mathbf{M}(z)}. \tag{34.66}$$

Пример 34.9

Предположим, что среднее нормального распределения известно и равно μ. Проверим гипотезу о дисперсии H_0: $\sigma^2 = \sigma_0^2$ против H_1: $\sigma^2 = \sigma_1^2$. Получаем

$$Z_m = \sum z_i = -m \log \sigma_1 - \\ - \frac{1}{2\sigma_1^2} \sum (x-\mu)^2 + m \log \sigma_0 + \frac{1}{2\sigma_0^2} \sum (x-\mu)^2. \quad (34.67)$$

Эта величина лежит между $\log\{\beta/(1-\alpha)\}$ и $\log\{(1-\beta)/\alpha\}$, если

$$\log \frac{\beta}{1-\alpha} < -m \log \frac{\sigma_1}{\sigma_0} - \frac{1}{2}\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2}\right) \sum (x-\mu)^2 < \log \frac{1-\beta}{\alpha}. \quad (34.68)$$

После некоторых преобразований находим, что это эквивалентно

$$\frac{2 \log \frac{1-\beta}{\alpha} + m \log \frac{\sigma_1^2}{\sigma_0^2}}{\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}} \leqslant \sum (x-\mu)^2 \leqslant \frac{2 \log \frac{\beta}{1-\alpha} + m \log \frac{\sigma_1^2}{\sigma_0^2}}{\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}}. \quad (34.69)$$

ОХ и СОВ даны в упражнениях 34.18 и 34.19.

Если среднее неизвестно, критерий остается тем же самым, за исключением того, что статистика критерия $\sum (x-\mu)^2$ заменяется на $\sum (x-\bar{x})^2$, а значение m в неравенстве (34.69) заменяется на $(m-1)$.

Эффективность последовательного критерия

34.22 В общем случае для заданных α и β, θ_0 и θ_1 может быть получено много различных критериев. Нет смысла сравнивать их по мощности для данных объемов выборок, потому что они устроены так, чтобы иметь одинаковую β-ошибку. Мы можем, однако, определить эффективность через объем выборки или СОВ. Критерий с меньшим СОВ разумно называть более эффективным. Следуя Вальду (1947), мы докажем, что если пренебречь граничными эффектами, то ПКОВ является наиболее эффективным. Точнее, если S' есть ПКОВ, а S — некоторый другой критерий, основанный на сумме логарифмов одинаково распределенных величин, то

$$\mathbf{M}_i(n \mid S) \geqslant \mathbf{M}_i(n \mid S'), \quad i = 0, 1, \quad (34.70)$$

где $\mathbf{M}_i$ обозначает среднее значение n при гипотезе H_i.

Заметим прежде всего, что если u — любая случайная величина, то $u - \mathbf{M}(u)$ есть ее отклонение от среднего и

$$\exp\{u - \mathbf{M}(u)\} \geqslant 1 + \{u - \mathbf{M}(u)\}.$$

Беря математические ожидания от обеих частей, получаем

$$\mathbf{M}[\exp\{u - \mathbf{M}(u)\}] \geqslant 1, \tag{34.71}$$

откуда

$$\mathbf{M}(\exp u) \geqslant \exp\{\mathbf{M}(u)\}. \tag{34.72}$$

Кроме того, для *любого* замкнутого последовательного критерия, основанного на суммах типа Z_n, из (34.42) имеем

$$\mathbf{M}_i(n \mid S) = \frac{\mathbf{M}_i(\log L_n \mid S)}{\mathbf{M}_i(z)}. \tag{34.73}$$

Пусть $\mathbf{M}^*$ обозначает математическое ожидание, когда верна H_0, а $\mathbf{M}^{**}$ — математическое ожидание, когда верна H_1. Пренебрегая, как в (34.22), граничными эффектами, находим

$$\mathbf{M}^*(L_n \mid S) = \frac{\beta}{1-\alpha} \tag{34.74}$$

и аналогично, как в (34.21),

$$\mathbf{M}^{**}(L_n \mid S) = \frac{1-\beta}{\alpha}. \tag{34.75}$$

Следовательно,

$$\mathbf{M}_0(n \mid S) = \frac{1}{\mathbf{M}_0(z)}\{(1-\alpha)\,\mathbf{M}^*(\log L_n \mid S) + \alpha \mathbf{M}^{**}(\log L_n \mid S)\}. \tag{34.76}$$

Тогда в силу (34.72), (34.74), (34.75) $\mathbf{M}_0(z) \leqslant 0$ и

$$\mathbf{M}_0(n \mid S) \geqslant \frac{1}{\mathbf{M}_0(z)}\left\{(1-\alpha)\log\frac{\beta}{1-\alpha} + \alpha\log\frac{1-\beta}{\alpha}\right\}, \tag{34.77}$$

а перестановка H_0 и H_1, α и β в (34.77) дает

$$\mathbf{M}_1(n \mid S) \geqslant \frac{1}{\mathbf{M}_1(z)}\left\{\beta\log\frac{\beta}{1-\alpha} + (1-\beta)\log\frac{1-\beta}{\alpha}\right\}. \tag{34.78}$$

Когда $S = S'$, эти неравенства, как в (34.43), заменяются (если пренебрегать граничными эффектами) равенствами. Таким образом, (34.70) доказано.

Пример 34.10

Одним из преимуществ последовательного метода, как мы отмечали, является то, что при данных (α, β) требуется в среднем меньшая выборка, чем для метода с фиксированным объемом. Трудно получить общие формулы сравнения двух методов, но можно проиллюстрировать указанный факт на задаче проверки гипотезы о среднем в нормальном распределении (пример 34.8).

При фиксированных n и α критерий состоит в нахождении отклонения d такого, что

$$\mathbf{P}\{\mu_0 - d \leqslant \bar{x} \leqslant \mu_0 + d \mid H_0\} = 1 - \alpha,$$
$$\mathbf{P}\{\mu_0 - d \leqslant \bar{x} \leqslant \mu_0 + d \mid H_1\} = \beta.$$

Полагая

$$\lambda_0 = \sqrt{n(d - \mu_0)}, \quad \lambda_1 = \sqrt{n(d - \mu_1)},$$

имеем

$$n = \frac{(\lambda_1 - \lambda_0)^2}{(\mu_0 - \mu_1)^2}. \tag{34.79}$$

По данным α, β и μ_1 определяется n. Сравним его с СОВ для ПКОВ. Беря приближенную формулу (34.43), которая записывается в виде

$$\mathbf{M}_i(n) = \frac{1}{\mathbf{M}_i(z)}[K(\mu)\log B + \{1 - K(\mu)\}\log A], \tag{34.80}$$

и учитывая, что

$$\mathbf{M}_1(z) = \frac{1}{2}(\mu_0 - \mu_1)^2 \quad \text{и} \quad \mathbf{M}_0(z) = -\frac{1}{2}(\mu_0 - \mu_1)^2,$$

находим

$$\frac{\mathbf{M}_1(n)}{n} = \frac{2}{(\lambda_1 - \lambda_0)^2}\{\beta \log B + (1 - \beta)\log A\}. \tag{34.81}$$

Аналогично получаем

$$\frac{\mathbf{M}_0(n)}{n} = -\frac{2}{(\lambda_1 - \lambda_0)^2}\{(1 - \alpha)\log B + \alpha \log A\}. \tag{34.82}$$

Таким образом, для $\alpha = 0{,}01$, $\beta = 0{,}03$ имеем $A = 97$, $B = 3/99$ и $\lambda_0 = 2{,}5758$, $\lambda_1 = -1{,}8808$. Отношение $\mathbf{M}_0(n)/n$ равно тогда 0,43, а $\mathbf{M}_1(n)/n = 0{,}55$. Таким образом, в последовательном случае нам требуется в среднем либо 43, либо 55% от выборки фиксированного объема, требуемой для достижения таких же ошибок.

Следует подчеркнуть, что сокращение объема выборки в случае ПКОВ, вообще говоря, обнаружится только тогда, когда верна одна из гипотез H_0 и H_1, и может не иметь места при других значениях параметра. Гарантированная экономия, следовательно, ограничивается случаями, когда конкурирующие гипотезы могут быть установлены заранее. См. Андерсон (1960).

Сложные гипотезы

34.23 Хотя мы рассматривали критерий для простой гипотезы H_0 против простой H_1, ОХ и СОВ могут быть вычислены на самом деле для всех возможных альтернативных значений и, следовательно, дадут нам характеристики критерия для простой

гипотезы H_0 против сложной H_1. Рассмотрим теперь случай сложной гипотезы H_0. Предположим, что θ может изменяться в некоторой области Ω. Нам нужно проверить гипотезу о том, что θ принадлежит некоторой ее подобласти ω_a (области принятия), против альтернатив, что θ лежит либо в области отклонения ω_r, либо в области безразличия $\Omega - \omega_a - \omega_r$ (которая может быть пустой). Мы будем требовать от ошибок двух свойств: вероятность ошибки первого рода $\alpha(\theta)$, которая в общем случае изменяется вместе с θ, не должна превышать фиксированного числа α для всех θ из ω_a, а вероятность ошибки второго рода $\beta(\theta)$ не должна превышать β для всех θ из ω_r. Где бы в действительности ни находился параметр θ, ошибки будут ограничены заданными α и β.

34.24 Такие условия, однако, вряд ли могут привести к полезным результатам. Мы имеем гарантию на все случаи, но излишняя гарантия может привести к значительной потере эффективности.

Вальд (1947) считает, что в качестве разумного требования, возможно, лучше рассматривать усредненные значения $\alpha(\theta)$ по ω_a и $\beta(\theta)$ по ω_r. Возникает вопрос, какого сорта среднее следует использовать. Вальд вводит две весовые функции $w_a(\theta)$ и $w_r(\theta)$ такие, что

$$\int_{\omega_a} w_a(\theta)\, d\theta = 1, \qquad \int_{\omega_r} w_r(\theta)\, d\theta = 1, \tag{34.83}$$

и затем полагает

$$\int_{\omega_a} w_a(\theta)\, \alpha(\theta)\, d\theta = \alpha, \tag{34.84}$$

$$\int_{\omega_r} w_r(\theta)\, \beta(\theta)\, d\theta = \beta. \tag{34.85}$$

С помощью этих средних задача сводится к задаче о проверке простой гипотезы. Действительно, если обозначить

$$L_{0m} = \int_{\omega_a} f(x_1, \theta) f(x_2, \theta) \ldots f(x_m, \theta)\, w_a(\theta)\, d\theta, \tag{34.86}$$

$$L_{1m} = \int_{\omega_r} f(x_1, \theta) f(x_2, \theta) \ldots f(x_m, \theta)\, w_r(\theta)\, d\theta, \tag{34.87}$$

то отношение правдоподобия L_{0m}/L_{1m} может быть использовано обычным способом с ошибками α и β. Мы можем, если угодно, рассматривать (34.86) и (34.87) как апостериорные вероятности

выборки, когда само θ имеет априорные вероятности $w_a(\theta)$ и $w_r(\theta)$.

34.25 Указанная процедура, разумеется, сводит задачу к задаче о нахождении или выборе весовых функций $w_a(\theta)$ и $w_r(\theta)$. Мы оказываемся в точно таком же положении, как когда хотим применить теорему Байеса. Мы можем применить одну из форм постулата Байеса, например, предположив, что $w_a(\theta) = 1$ всюду в ω_a. Другая возможность для выбора $w_a(\theta)$ и $w_r(\theta)$ заключается в том, чтобы оптимизировать некоторые свойства критерия.

Например, выбор критерия сделан, когда выбраны α и β (или, при нашей аппроксимации, A и B) и весовые функции. Среди всех таких критериев будут находиться критерии с максимальными значениями $\alpha(\theta)$ и $\beta(\theta)$. Если мы выберем весовые функции так, чтобы минимизировать $(\max \alpha, \max \beta)$, то получим критерий, который для данных A и B имеет наименьшую возможную границу средних ошибок. Если невозможно минимизировать максимумы α и β одновременно, то можно попытаться минимизировать максимум некоторой функции от них, например, от суммы.

Последовательный t-критерий

34.26 Критерий, предложенный Вальдом (1947) и в модифицированном виде другими авторами, служит для проверки гипотезы о среднем значении μ нормального распределения, когда дисперсия неизвестна. Он называется последовательным t-критерием, потому что имеет дело с такой же задачей, как и t-критерий Стьюдента в случае фиксированной выборки; но исключение зависимости от масштабного параметра σ в нем происходит по-иному и его название, возможно, несколько вводит в заблуждение.

Говоря точнее, мы хотим проверить гипотезу о том, что по сравнению с некоторым значением μ_0 отклонение $(\mu - \mu_0)/\sigma$ мало, скажем $< \delta$. Тогда три подобласти из **34.23** имеют следующий вид:

ω_a состоит из значений (μ_0, σ), σ — любое;

ω_r состоит из значений, для которых $|\mu - \mu_0| \geqslant \sigma\delta$, σ — любое;

$\Omega - \omega_a - \omega_r$ состоит из значений, для которых $0 \leqslant |\mu - \mu_0| < \sigma\delta$, σ — любое.

Определим следующим образом весовые функции для σ:

$$v_{ac} = \begin{cases} 1/c, & 0 \leqslant \sigma \leqslant c, \\ 0 & \text{в остальных случаях,} \end{cases} \tag{34.88}$$

$$v_{rc} = \begin{cases} 1/2c, & 0 \leqslant \sigma \leqslant c, \quad \mu = \mu_0 + \delta\sigma, \\ 0 & \text{в остальных случаях.} \end{cases} \tag{34.89}$$

Тогда

$$L_{1m} = \int v_{rc} \frac{1}{(2\pi)^{m/2}\,\sigma^m} \exp\left\{-\frac{1}{2\sigma^2}\sum (x_i-\mu)^2\right\} d\sigma =$$

$$= \frac{1}{(2\pi)^{m/2}} \frac{1}{2c} \int_0^c \left[\frac{1}{\sigma^m} \exp\left\{-\frac{1}{2\sigma^2}\sum (x_i-\mu_0-\delta\sigma)^2\right\} + \right.$$

$$\left. + \frac{1}{\sigma^m} \exp\left\{-\frac{1}{2\sigma^2}\sum (x_i-\mu_0+\delta\sigma)^2\right\}\right] d\sigma, \qquad (34.90)$$

$$L_{0m} = \frac{1}{(2\pi)^{m/2}} \frac{1}{c} \int_0^c \frac{1}{\sigma^m} \exp\left\{-\frac{1}{2\sigma^2}\sum (x_i-\mu_0)^2\right\} d\sigma. \qquad (34.91)$$

Когда c стремится к бесконечности, предел отношения L_{1m}/L_{0m} равен

$$\lim \frac{L_{1m}}{L_{0m}} = \frac{\frac{1}{2}\int_0^\infty \frac{d\sigma}{\sigma^m}\left[\exp\left\{-\frac{1}{2\sigma^2}\sum (x_i-\mu_0-\delta\sigma)^2\right\} + \exp\left\{-\frac{1}{2\sigma^2}\sum (x_i-\mu_0+\delta\sigma)^2\right\}\right]}{\int_0^\infty \frac{d\sigma}{\sigma^m} \exp\left\{-\frac{1}{2\sigma^2}\sum (x_i-\mu_0)^2\right\}}. \qquad (34.92)$$

Он зависит от наблюдаемых иксов, а также от заданных μ_0 и δ, но не зависит от параметра σ, который исключается в результате интегрирования с весовыми функциями (34.88) и (34.89). Если мы сможем вычислить интегралы в (34.92), то сможем использовать это отношение для получения последовательного критерия.

34.27 Весовые функции (34.88) и (34.89), выбранные на первый взгляд достаточно произвольно, на самом деле оптимизируют критерий. Чтобы это доказать, мы прежде всего установим, (а) что $\alpha(\mu, \sigma)$ постоянна в ω_a, (б) что $\beta(\mu, \sigma)$ зависит только от $|(\mu-\mu_0)/\sigma|$ и (в) что $\beta(\mu, \sigma)$ монотонно убывает относительно $|(\mu-\mu_0)/\sigma|$.

Если $\bar{x}$ — выборочное среднее, а $S^2 = \sum (x_i-\bar{x})^2$, то распределение отношения $(\bar{x}-\mu)/S$ зависит только от $(\mu-\mu_0)/\sigma$. Если мы сможем показать, что (34.92) представляет собой однозначную функцию от $(\bar{x}-\mu_0)/S$, то свойства (а) и (б) будут следовать немедленно, потому что $(\mu-\mu_0)/\sigma$ равно нулю в ω_a, а $\beta(\mu, \sigma)$ зависит только от распределения (34.92). Числитель и знаменатель в (34.92) являются однородными функциями от $(x_i-\mu_0)$ степени $m-1$, что можно проверить, подставляя $x_i = \lambda x_i$, $\mu_0 = \lambda\mu_0$, $\sigma = \lambda\sigma$. Таким образом, отношение (34.92) имеет нулевую степень. Кроме того, оно является функцией

только от $\sum (x-\mu_0)^2$ и $\sum (x-\mu_0)$, и, следовательно, мы не изменим его, подставляя $(x_i-\mu_0)/\sqrt{\sum (x_i-\mu_0)^2}$ вместо $x_i-\mu_0$. Тогда данное отношение будет функцией только от $\sum (x_i-\mu_0)/\sqrt{\sum (x_i-\mu_0)^2}$ и в действительности функцией от квадрата этой величины, а именно от

$$\frac{(\bar{x}-\mu_0)^2}{\sum (x_i-\mu_0)^2} = \frac{(\bar{x}-\mu_0)^2}{n(\bar{x}-\mu_0)^2+S^2}.$$

Оно является, следовательно, однозначной функцией от $(\bar{x}-\mu_0)/S$.

Чтобы показать, что $\beta(\mu, \sigma)$ монотонно убывает по $|(\mu-\mu_0)/\sigma|$, достаточно показать, что отношение (34.92) является строго возрастающей функцией от $|(\bar{x}-\mu_0)/S|$ или, что эквивалентно, от $(\bar{x}-\mu_0)^2/\sum (x_i-\mu_0)^2$. При фиксированной $\sum (x_i-\mu_0)^2$ знаменатель в (34.92) фиксирован, а числитель является возрастающей функцией от $(\bar{x}-\mu_0)^2$. Таким образом, все отношение возрастает по $(\bar{x}-\mu_0)^2$ при фиксированной $\sum (x_i-\mu_0)^2$, и отсюда следует требуемый результат.

34.28 При этих условиях мы можем доказать, что последовательный t-критерий оптимален. В самом деле, критерий оптимален, если (i) $\alpha(\theta)$ постоянна в ω_a; (ii) $\beta(\theta)$ постоянна на границе ω_r; (iii) $\beta(\theta)$ при любом θ внутри ω_r не превосходит своих граничных значений.

Чтобы доказать это, предположим, что v_a и v_r — весовые функции, удовлетворяющие данным условиям, а w_a, w_r — две другие весовые функции. Пусть α, β — ошибки для первого набора весовых функций, α^*, β^* — для второго. Тогда имеем

$$\frac{1-\beta}{\alpha} = A, \qquad \frac{\beta}{1-\alpha} = B$$

и, следовательно,

$$\int_{\omega_a} \alpha^*(\theta)\, w_a(\theta)\, d\theta = \alpha = \frac{1-B}{A-B}, \tag{34.93}$$

$$\int_{\omega_r} \beta^*(\theta)\, w_r(\theta)\, d\theta = \beta = \frac{B(A-1)}{A-B}. \tag{34.94}$$

Таким образом, в ω_a максимум $\alpha^*(\theta)$ больше, чем $(1-B)/(A-B)$, потому что интеграл от $w_a(\theta)$ по этой области равен единице. Но если ошибка α, отвечающая v_a, постоянна в этой области, то ее максимум равен $(1-B)/(A-B)$. Следовательно,

$$\max \alpha^*(\theta) \geqslant \max \alpha(\theta) \quad \text{в} \quad \omega_a.$$

Аналогично для ω_r. Максимум $\beta^*(\theta)$ достигается вне ω_r и не может превосходить $B(A-1)/(A-B)$, тогда как для $\beta(\theta)$ максимальное значение должно быть где-то достигнуто. Следовательно, $\max \beta^*(\theta) \geqslant \max \beta(\theta)$ в ω_r. Утверждение доказано. Условия, которые мы рассмотрели, достаточны, но отнюдь не необходимы.

34.29 Некоторые таблицы для пользования этим критерием были даны Армитеджем (1947). Интегралы, входящие в (34.92), выражаются через вырожденные гипергеометрические функции и, следовательно, через нецентральное t-распределение. *Tables to facilitate sequential t-Tests* (U. S. A. National Bureau of Standards, Applied Mathematics Series, 7, 1951) являются совместной работой нескольких авторов с введением К. Арнольда. См. также более поздние работы Армитеджа и других и недавнюю работу Майерса и др. (1966).

Иной метод решения предложен Д. Коксом (1952a, b) в случае, когда распределение системы достаточных статистик факторизуется, как в (23.118). Тогда обычным способом ПКОВ для θ_l можно основать на статистике T_s, не зависящей от мешающего (часто масштабного) параметра θ_k. Данный подход, кроме того, требует условия инвариантности (см. Холл и др. (1965)). В этом же направлении Хэджнэл (1961) развил последовательный t-критерий, основанный на двойной выборке.

В случае большой выборки Д. Кокс (1963) предложил последовательный критерий для любой сложной гипотезы, основанный на теории МП (см. упражнение 34.21).

Последовательное оценивание

34.30 При проверке гипотез мы обычно фиксируем ошибки заранее и затем производим выборку до тех пор, пока не примем или не отвергнем гипотезу. Можно также применить последовательный процесс к оцениванию, но тогда наши задачи решаются труднее, а возможно, даже труднее формулируются. Мы проводим последовательность наблюдений с целью оценить некоторый параметр исходного распределения, но в общем случае нелегко различить, какая оценка является подходящей, каковы смещения, каковы выборочные ошибки или какими должны быть правила, определяющие окончание выборочного процесса. Основная трудность состоит в том, что объем выборки является случайной величиной. Второй неприятностью является граничный эффект, о котором мы уже говорили *).

*) Леман и Стейн (1950) рассматривали понятие «полноты» в последовательном случае, но общие критерии трудно применять даже к выборке по качественному признаку (см. Де Гроот (1959)).

34.31 Полученный нами в (34.42) результат эквивалентен тому, что

$$\mathbf{M}\{Z_n - n\mathbf{M}(z)\} = 0. \tag{34.95}$$

Предположим, что существуют вторые абсолютные моменты, что дисперсия каждого z_i равна σ^2 и что существует $\mathbf{M}(n^2)$. Тогда дисперсия $\{Z_n - n\mathbf{M}(z)\}$ равна (доказательство предоставлено читателю в качестве упражнения 34.5)

$$\mathbf{M}\{Z_n - n\mathbf{M}(z)\}^2 = \sigma^2\mathbf{M}(n), \tag{34.96}$$

и, следовательно, полагая, как и прежде, $\mathbf{M}(z) = \mu$, находим

$$\mu^2\mathbf{M}(n^2) = \sigma^2\mathbf{M}(n) + 2\mu\mathbf{M}(nZ_n) - \mathbf{M}(Z_n^2).$$

Если Z_n и n некоррелированы, то это приводится к

$$\mu^2\mathbf{D}n = \sigma^2\mathbf{M}(n) - \mathbf{D}Z_n. \tag{34.97}$$

Для более высоких моментов результаты были получены Волфовицем (1947) (см. упражнение 34.6). См. также Чоу и др. (1965).

34.32 Пусть теперь

$$Y_n = \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log f(x_i, \theta). \tag{34.98}$$

Так как при условиях регулярности $\mathbf{M}\left(\frac{\partial \log f}{\partial\theta}\right) = 0$, то на основании соотношений типа (34.95) находим

$$\mathbf{M}(Y_n) = 0, \tag{34.99}$$

и

$$\mathbf{D}Y_n = \mathbf{M}\left(\sum \frac{\partial \log f}{\partial\theta}\right)^2 = \mathbf{M}(n)\,\mathbf{M}\left(\frac{\partial \log f}{\partial\theta}\right)^2. \tag{34.100}$$

Если t — оценка параметра θ, имеющая смещение $b(\theta)$, т. е. $\mathbf{M}(t) = \theta + b(\theta)$, то, дифференцируя это равенство, получаем

$$\operatorname{cov}\left(t, \sum \frac{\partial \log f}{\partial\theta}\right) = \mathbf{M}\left(t \sum \frac{\partial \log f}{\partial\theta}\right) = 1 + b'(\theta). \tag{34.101}$$

Далее, в силу неравенства Коши — Буняковского

$$\mathbf{D}t\,\mathbf{M}\left(\sum \frac{\partial \log f}{\partial\theta}\right)^2 \geqslant \{1 + b'(\theta)\}^2,$$

и, следовательно, согласно (34.100)

$$\mathbf{D}t \geqslant \frac{\{1 + b'(\theta)\}^2}{\mathbf{M}(n)\,\mathbf{M}\left(\frac{\partial \log f}{\partial\theta}\right)^2}. \tag{34.102}$$

Это нижняя граница дисперсии при последовательном оценивании, полученная Волфовицем. Она представляет собой МГД (17.22) с $\mathbf{M}(n)$ вместо n. Волфовиц (1947) распространяет этот результат также на совместную оценку нескольких параметров.

Пример 34.11

Рассмотрим биномиальное распределение

$$f(x, \omega) = \omega^x (1-\omega)^{1-x}, \quad x = 0, 1.$$

Имеем

$$\frac{\partial \log f}{\partial \omega} = \frac{x}{\omega} - \frac{1-x}{1-\omega}, \quad \mathbf{M}\left(\frac{\partial \log f}{\partial \omega}\right)^2 = \frac{1}{\omega(1-\omega)}.$$

Если p — несмещенная оценка для ω по выборке из этого распределения, то

$$\mathbf{D}p \geqslant \frac{\omega(1-\omega)}{\mathbf{M}(n)}. \tag{34.103}$$

34.33 При большом n теория последовательного оценивания сильно упрощается благодаря общему результату, принадлежащему Энскомбу (1949b, 1952, 1953). Если этот результат кратко сформулировать, то он звучит так: у статистик, для которых имеет место центральная предельная теорема, формулы стандартных ошибок в случае последовательных выборок такие же, как и в случае выборок фиксированного объема. Эвристически это можно вывести из (34.102). Величина n изменяется относительно своего среднего n_0 со стандартным отклонением порядка $n_0^{-1/2}$, и, таким образом, формулы, имеющие точность порядка n^{-1}, остаются столь же точными, если мы используем n_0 вместо n. Более формально:

Пусть $\{Y_n\}$, $n = 1, 2, \ldots$, — последовательность случайных величин. Пусть существуют действительное число θ, последовательность положительных чисел $\{w_n\}$ и функция распределения $F(x)$ такие, что

(а) Y_n сходятся к θ в масштабе w_n, а именно:

$$\mathbf{P}\left\{\frac{Y_n - \theta}{w_n} \leqslant x\right\} \to F(x), \quad \text{когда} \quad n \to \infty; \tag{34.104}$$

(б) $\{Y_n\}$ равномерно непрерывны по вероятности, а именно: для данных (малых) положительных ε и η

$$\mathbf{P}\left\{\left|\frac{Y_{n'} - Y_n}{w_n}\right| < \varepsilon \text{ для всех } n, n' \text{ таких, что } |n' - n| < \varepsilon n\right\} > 1 - \eta. \tag{34.105}$$

Пусть $\{n_r\}$ — возрастающая последовательность положительных целых чисел, стремящаяся к бесконечности, и $\{N_r\}$ — последовательность случайных величин, принимающих целые положительные значения, такая, что $N_r/n_r \to 1$ по вероятности при $r \to \infty$. Тогда

$$\mathbf{P}\left\{\frac{Y_{N_r}-\theta}{w_{N_r}} \leqslant x\right\} \to F(x), \quad \text{когда} \quad r \to \infty, \tag{34.106}$$

во всех точках непрерывности $F(x)$.

Сложность формулировок и доказательств обусловлена особенностями, на которые мы уже обращали внимание: граничные эффекты (представленные соотношением между N_r и n_r) и изменение n_r.

Итак, пусть выполняется (34.105); возьмем достаточно большое ν такое, что для любого $n_r > \nu$

$$\mathbf{P}\{|N_r - n_r| < cn_r\} > 1 - \eta. \tag{34.107}$$

Рассмотрим события:

E: $|N_r - n_r| < cn_r$ и $|Y_{N_r} - Y_{n_r}| < \varepsilon w_{N_r}$,

A: $|Y_{n'} - Y_n| < \varepsilon w_n$ при всех n' таких, что $|n' - n| < \varepsilon n$,

B: $|N_r - n_r| < cn_r$.

Тогда

$$\mathbf{P}(E) \geqslant \mathbf{P}\{A \text{ и } B\} = \mathbf{P}(A) - \mathbf{P}\{A \text{ и не-}B\} \geqslant$$
$$\geqslant \mathbf{P}(A) - \mathbf{P}(\text{не-}B) \geqslant 1 - 2\eta. \tag{34.108}$$

Также

$$\mathbf{P}\{Y_{N_r} - \theta \leqslant x w_{N_r}\} = \mathbf{P}\{Y_{N_r} - \theta \leqslant x w_{n_r} \text{ и } E\} +$$
$$+ \mathbf{P}\{Y_{N_r} - \theta \leqslant x w_{n_r} \text{ и не-}E\}.$$

Таким образом, используя определение события E, находим

$$\mathbf{P}\{Y_{n_r} - \theta \leqslant (x-\varepsilon) w_{n_r}\} - 2\eta < \mathbf{P}\{Y_{n_r} - \theta \leqslant x w_{n_r}\} <$$
$$< \mathbf{P}\{Y_{n_r} - \theta \leqslant (x+\varepsilon) w_{n_r}\} + 2\eta,$$

откуда следует (34.106). Необходимо заметить, что доказательство проведено без предположения о независимости N_r и Y_n.

34.34 Применим этот результат к последовательному оцениванию. Пусть $x_1, x_2, \ldots$ — последовательность наблюдений, Y_n — оценка параметра θ, D_n — оценка масштабного параметра w_n для Y_n. Выборочное правило таково: задаем некоторую постоянную k и производим выбор до первого осуществления неравенства $D_n \leqslant k$, а затем вычисляем Y_n. Покажем, что Y_n является

оценкой θ с масштабным параметром, асимптотически равным k, если k мало.

Пусть выполнены условия (34.104) и (34.105) и $\{k_r\}$ — последовательность положительных чисел, сходящаяся к нулю. Пусть $\{N_r\}$ — последовательность случайных величин таких, что N_r равно наименьшему целому n, для которого $D_n \leqslant k_r$, и пусть $\{n_r\}$ — последовательность чисел таких, что n_r равно наименьшему n, при котором $w_n \leqslant k_r$. Мы потребуем, чтобы выполнялись два дополнительных условия:

(в) $\{w_n\}$ монотонно сходится к нулю и $w_n/w_{n+1} \to 1$, когда $n \to \infty$;

(г) N_r является случайной величиной для всех r и $N_r/n_r \to 1$ по вероятности, когда $r \to \infty$.

Условие (в) влечет $w_{n_r}/k_r \to 1$ при $n \to \infty$. Из нашего предыдущего результата тогда следует, что

$$\mathbf{P}\left\{\frac{Y_{N_r} - \theta}{k_r} \leqslant x\right\} \to F(x), \quad \text{когда} \quad r \to \infty. \tag{34.109}$$

34.35 Можно также показать, что если иксы независимы и одинаково распределены, то из условий (а) и (в) (которые легко проверяются) вытекает условие (б) и распределение суммы иксов сходится к некоторой функции распределения. В частности, эти условия выполняются для оценок максимального правдоподобия, для оценок, основанных на средних значениях некоторых функций от наблюдений, и для квантилей.

Пример 34.12

Рассмотрим оценивание среднего значения μ нормального распределения с неизвестной дисперсией σ^2. Нам требуется оценка с (малой) дисперсией k^2.

Очевидной статистикой служит $Y_n = \bar{x}_n$. Для фиксированного n ее дисперсия равна σ^2/n и оценивается через

$$D_n^2 = \frac{1}{n(n-1)} \sum (x_i - \bar{x})^2. \tag{34.110}$$

Условия (а) и (в), очевидно, выполняются, и благодаря результату, указанному в **34.35**, это влечет за собой выполнение условия (б). Чтобы показать, что выполняется (г), применим преобразование Хельмерта

$$\xi_i = \left(x_{i+1} - \frac{1}{i}\sum_{j=1}^{i} x_j\right)\sqrt{\frac{i}{i+1}}.$$

Тогда

$$D_n^2 = \frac{1}{n(n-1)} \sum_{i=1}^{n-1} \xi_i^2.$$

Согласно усиленному закону больших чисел при заданных ε, η найдется ν такое, что

$$\mathrm{P}\left\{\left|\frac{1}{n-1}\sum_{i=1}^{n-1}\xi_i^2-\sigma^2\right|<\varepsilon \text{ для всех } n>\nu\right\}>1-\eta. \qquad (34.111)$$

Если k достаточно мало, то вероятность того, что $D_n \geqslant k$ для любого n в интервале $2 \leqslant n \leqslant \nu$, превосходит $1-\eta$. Таким образом, при условии, что $N > \nu$, из (34.111) следует, что

$$\left|\frac{N}{\sigma^2/k^2}-1\right|<\frac{\varepsilon}{\sigma^2}$$

с вероятностью, превосходящей $1-\eta$. Следовательно, когда k стремится к нулю, условие (г) выполняется.

Здесь правило состоит в том, что мы задаем k и производим выбор, пока не окажется $D_n \leqslant k$. Тогда дисперсия среднего значения $\bar{x}$ равна приблизительно k^2.

Пример 34.13

Рассмотрим пуассоновское распределение с параметром λ. Если мы продолжаем выбор, пока дисперсия среднего, оцениваемая как $\bar{x}/n$, не станет меньше k^2, то для λ мы получаем оценку $\bar{x}$ с дисперсией k^2. Это эквивалентно продолжению выбора до тех пор, пока число успехов не станет меньше k^2n^2. Но не следует использовать этот результат при малых n.

С другой стороны, предположим, что мы хотим задавать заранее не дисперсию, а коэффициент вариации l. Тогда указанный метод не будет работать. Он означал бы, что мы продолжаем выбор, пока $\bar{x}/\sqrt{\bar{x}/n}$ не станет меньше, чем l, т. е. $n\bar{x} \leqslant l^2$, или пока сумма наблюдений не станет меньше l^2. Но сумма должна в конце концов превзойти любое конечное число. Это связано с результатом, отмеченным в примере 34.1, где мы видели, что в случае последовательного выбора редких качественных признаков коэффициент вариации приблизительно постоянен.

Метод двойной выборки Стейна

34.36 В конце примера 23.7 мы видели, что при фиксированном n не существует подобного критерия для среднего значения μ нормальной совокупности с неизвестной дисперсией σ^2, который имел бы мощность, не зависящую от σ^2. Отсюда следует (см. **23.26**), что для μ нельзя найти доверительного интервала заранее заданной длины. Однако, как указал Стейн (1945), эти утверждения несправедливы, если мы используем последовательный метод.

34.37 Рассмотрим нормальную совокупность со средним μ и дисперсией σ^2. Пусть нам требуется построить для μ доверительный интервал длины l с коэффициентом доверия $1-\alpha$. Мы сделаем сначала выборку фиксированного объема n_0, а затем выборку объема $n-n_0$, где n теперь зависит от наблюдений в первой выборке.

Возьмем t-статистику Стьюдента с n_0-1 степенями свободы, и пусть вероятность того, что она лежит в интервале от $-t_\alpha$ до t_α, равна $1-\alpha$. Положим

$$\sqrt{z}=\frac{l}{2t_\alpha}. \tag{34.112}$$

Пусть s^2 — оценка дисперсии по выборке объема n_0, т. е.

$$s^2=\frac{1}{n_0-1}\sum_{i=1}^{n_0}(x_i-\bar{x})^2. \tag{34.113}$$

Определим n, полагая

$$n=\max\{n_0,\ 1+[s^2/z]\}, \tag{34.114}$$

где $[s^2/z]$ означает целую часть числа s^2/z.

Рассмотрим все вместе n наблюдений, и пусть их среднее значение равно Y_n. Тогда Y_n не зависит от s и, следовательно, $(Y_n-\mu)\sqrt{n}$ не зависит от s, откуда вытекает, что величина $(Y_n-\mu)\sqrt{n}/s$ распределена как t с n_0-1 степенями свободы. Таким образом,

$$\mathsf{P}\left\{\left|\frac{(Y_n-\mu)\sqrt{n}}{s}\right|<t_\alpha\right\}=1-\alpha,$$

или

$$\mathsf{P}\left\{Y_n-\frac{st_\alpha}{\sqrt{n}}\leqslant\mu\leqslant Y_n+\frac{st_\alpha}{\sqrt{n}}\right\}=1-\alpha,$$

или

$$\mathsf{P}\left\{Y_n-\frac{1}{2}l\leqslant\mu\leqslant Y_n+\frac{1}{2}l\right\}\geqslant 1-\alpha. \tag{34.115}$$

Появление неравенства в (34.115) обусловлено граничным эффектом, так как s^2/z может не быть целым. Граничный эффект, вообще говоря, мал, поэтому границы $Y_n\pm\frac{1}{2}l$ близки к точным границам для доверительного коэффициента $1-\alpha$. Фактически мы можем с помощью приема, предложенного Стейном, получить точные границы, хотя эта процедура связана с отбрасыванием наблюдений и, возможно, не представляет ценности для практики.

Силбайндер (1953) и Мошмэн (1958) рассматривали оптимальный выбор объема первой выборки в методе Стейна. Бхаттачарджи (1965) показал, что процедура Стейна более чувствительна к отклонению от нормальности, чем t-критерий Стьюдента (см. **31.3**), и что, как можно было ожидать, отклонение от нормальности вновь влечет зависимость длины интервала (и соответственно мощности критерия) от σ^2.

34.38 Чепмэн (1950) обобщил метод Стейна на критерий для отношения средних значений двух нормальных величин. Этот критерий не зависит от обеих дисперсий. Однако он зависит от распределения разности двух величин t, для которого Чепмэн приводит некоторые таблицы. Д. Кокс (1952c) занимался задачей оценивания по двойной выборке и получил ряд асимптотических результатов. Он также рассмотрел поправки к результатам, относящимся к одинарным и двойным выборкам, улучшающие приближение, даваемое асимптотической теорией. А. Бирнбаум и Хили (1960) обсуждали общий класс процедур при двойной выборке, позволяющих достигнуть заданной дисперсии. Грейбилл и Коннелл (1964) приводят двойную выборочную процедуру для оценивания дисперсии нормального закона внутри фиксированного интервала. Голдмэн и Цайглер (1966) сравнивали различные методы оценивания среднего или дисперсии нормального распределения: для среднего метод Стейна оказался наилучшим.

Критерии, свободные от распределения

34.39 Используя порядковые статистики, мы можем свести многие процедуры к биномиальному случаю. Рассмотрим, например, проверку гипотезы о том, что среднее нормального распределения больше, чем μ_0 (односторонний критерий). Заменим среднее медианой, а каждое наблюденное значение знаком $+$, если оно больше μ_0, и знаком $-$ в противоположном случае. При гипотезе H_0: $\mu = \mu_0$ эти знаки будут иметь биномиальное распределение с $\omega = 1/2$. При гипотезе H_1: $\mu = \mu_0 + k\sigma$ вероятность положительного знака равна

$$\omega_1 = \frac{1}{\sqrt{2\pi}} \int_{-k}^{\infty} \exp\left(-\frac{1}{2}x^2\right) dx. \qquad (34.116)$$

Обычным способом теперь можно применить ПКОВ для ω_0 против ω_1. Этот критерий будет иметь ошибку первого рода α и ошибку второго рода β от принятия H_0, когда верна H_1; ошибка второго рода будет $\leqslant\beta$, когда $\mu - \mu_0 > k\sigma$. Описанный метод является фактически последовательной формой критерия знаков из **32.2—7**.

Критерии такого сорта часто обладают значительной эффективностью, а потеря эффективности вполне может окупаться простотой применения. Армитедж (1947) сравнил приведенный выше критерий с t-критерием Вальда и пришел к выводу, что в отношении объема выборки оптимальный критерий несущественно лучше критерия знаков.

34.40 Джэксон (1960) привел библиографию последовательного анализа, классифицированную по темам. Джонсон (1961) дает полезный обзор этого предмета.

Решающие функции

34.41 В заключение этой главы мы можем кратко остановиться на развитии идей Вальда о последовательных процедурах в общую теорию решений. Пусть мы находимся в ситуации, когда, по крайней мере на некотором этапе выбора, мы должны принять решение, например принять гипотезу или продолжать выбор. Последствия этих решений предполагаются известными, и к тому же предполагается, что они могут быть выражены численно. Задача тогда состоит в выборе оптимальных решающих правил. Могут применяться различные возможные принципы, например, действовать так, чтобы максимизировать средний выигрыш или минимизировать средний проигрыш. Некоторые авторы доходят до утверждения, что все оценивание и проверка гипотез являются в действительности операциями принятия решения. Мы решительно не согласны как с тем, что все статистические задачи сводятся к принятию решений, так и с тем, что последствия многих решений можно выразить численно. И даже в тех случаях, когда можно принять оба эти положения, некоторые из предложенных принципов вызывают сомнения в том, что разумный человек стал бы использовать их на практике. То, что статистика представляет собой исключительно науку о принятии решений, кажется нам явным преувеличением. Но этот вопрос, подобно некоторым вопросам теории вероятностей, таков, что каждый должен составить о нем свое собственное мнение, пользуясь той помощью теории решающих функций, какую он может получить.

Основными руководствами по данной теории являются основополагающая работа самого Вальда (1950) и более поздняя книга Блекуэлла и Гиршика (1954).

УПРАЖНЕНИЯ

34.1 В условиях примера 34.1, используя упражнение 9.13, показать, что (34.3) влечет смещенность оценки m/n параметра ω.

34.2 Обращаясь к примеру 34.6, сделать набросок кривой OX для биномиального распределения с $\alpha = 0{,}01$, $\beta = 0{,}03$, $\omega_1 = 0{,}1$, $\omega_2 = 0{,}2$. (Эта кривая является половиной колоколообразной кривой с максимумом в $\omega = 0$ и нулем в $\omega = 1$. Шести точек достаточно, чтобы дать ее общий вид.) Аналогично, сделать набросок кривой COB для того же биномиального распределения.

34.3 Две выборки, каждая объема n, получены из совокупностей P_1 и P_2, в которых ω_1 и ω_2 — доли некоторого качественного признака. Из них образованы пары в порядке появления наблюдений. Пусть t_1 — число пар, в

которых присутствуют успех из P_1 и неудача из P_2; t_2 — число пар, в которых присутствуют неудача из P_1 и успех из P_2. Показать, что в (условном) множестве таких пар вероятность пары из t_1 равна

$$\omega = (1-\omega_1)\,\omega_2/\{\omega_1(1-\omega_2)+\omega_2(1-\omega_1)\}.$$

Рассматривая это как обычное биномиальное распределение на множестве $t = t_1 + t_2$ значений, показать, как проверить гипотезу $\omega_1 \geqslant \omega_2$, проверяя $\omega = 1/2$. Вывести отсюда последовательный критерий для $\omega_1 \geqslant \omega_2$.

Показать, что если

$$u = \frac{\omega_2(1-\omega_1)}{\omega_1(1-\omega_2)},$$

то $\omega = u/(1+u)$, и получить отсюда следующие приемочные и браковочные числа:

$$a_t = \frac{\log\dfrac{\beta}{1-\alpha}}{\log u_1 - \log u_0} + t\,\frac{\log\dfrac{1+u_1}{1+u_0}}{\log u_1 - \log u_0},$$

$$r_t = \frac{\log\dfrac{1-\beta}{\alpha}}{\log u_1 - \log u_0} + t\,\frac{\log\dfrac{1+u_1}{1+u_0}}{\log u_1 - \log u_0},$$

где u_i — значение u, соответствующее H_i $(i = 0, 1)$.

(Вальд, 1947.)

34.4 Пусть z — случайная величина такая, что $\mathbf{M}(z)$ существует и не равно нулю; пусть существует положительное δ такое, что $\mathbf{P}(e^z < 1-\delta) > 0$ и $\mathbf{P}(e^z > 1+\delta) > 0$, и пусть для любого действительного h существует $\mathbf{M}(\exp hz) = g(h)$. Для функции $h \neq 0$ из **34.14** показать тогда, что

$$\lim_{h\to\infty} g(h) = \infty = \lim_{h\to-\infty} g(h)$$

и что $g''(h) > 0$ для всех действительных значений h. Вывести отсюда, что $g(h)$ строго убывает на интервале $(-\infty, h^*)$ и строго возрастает на (h^*, ∞), где h^* — значение, при котором $g(h)$ минимальна. Как следствие показать, что существует не более одного значения $h \neq 0$, для которого $\mathbf{M}(\exp hz) = 1$.

(Вальд, 1947.)

34.5 В **34.31** вывести выражения (34.96), (34.97).

(См. Джонсон, 1959b.)

34.6 В условиях упражнения 34.5 показать, что третий момент величины $Z_n - n\mu$ равен

$$\mathbf{M}(Z_n - n\mu)^3 = \mu_3\mathbf{M}(n) - 3\sigma^2\mathbf{M}\{n(Z_n - n\mu)\},$$

где μ_3 — третий момент z.

(Волфовиц, 1947.)

34.7 Пусть z определено, как в (34.19), и t — комплексная переменная такая, что $\mathbf{M}(\exp zt) = \varphi(t)$ существует в некоторой части комплексной плоскости. Показать, что

$$\mathbf{M}\left[\{\exp(tZ_n)\}\{\varphi(t)\}^{-n}\right] = 1$$

для любой точки, в которой $|\varphi(t)| \geqslant 1$.

(Вальд, 1947.)

34.8 Положив в предыдущем упражнении $t = h$, доказать, что если $\mathbf{M}_b$ обозначает математическое ожидание при ограничении $Z_n \leqslant -b$ и $\mathbf{M}_a$ — при ограничении $Z_n \geqslant a$, то

$$K(h)\,\mathbf{M}_b \exp(hZ_n) + \{1 - K(h)\}\,\mathbf{M}_a \exp(hZ_n) = 1,$$

где K — оперативная характеристика. Отсюда, пренебрегая граничными эффектами, вывести, что

$$K(h) = \begin{cases} \dfrac{e^{h(a+b)} - e^{hb}}{e^{h(a+b)} - 1}, & h \neq 0, \\ \dfrac{a}{a+b}, & h = 0. \end{cases}$$

(Гиршик, 1946.)

34.9 Дифференцируя тождество из упражнения 34.7 по t и полагая $t = 0$, доказать что

$$\mathbf{M}(n) = \frac{a\{1 - K(h)\} - bK(h)}{\mathbf{M}(z)}.$$

Вывести отсюда равенство (34.43).

(Гиршик, 1946.)

34.10 Предполагая, как в предыдущем упражнении, что тождество дифференцируемо, получить результаты упражнений 34.7 и 34.8.

34.11 В тождестве упражнения 34.7 положить

$$-\log \varphi(t) = \tau,$$

где τ — чисто мнимое. Показать, что если функция $\varphi(t)$ не сингулярна при $t = 0$ и $t = h$, то это уравнение имеет два корня $t_1(\tau)$ и $t_2(\tau)$ для достаточно малых значений τ. Следуя упражнению 34.8, доказать, что характеристическая функция величины n асимптотически дается выражением

$$\mathbf{M}(e^{n\tau}) = \frac{A^{t_2} - A^{t_1} + B^{t_1} - B^{t_2}}{B^{t_1}A^{t_2} - A^{t_1}B^{t_2}}.$$

(Вальд, 1947.)

34.12 В случае, когда z — нормальная случайная величина со средним μ и дисперсией σ^2, доказать, что в упражнении 34.11 t_1 и t_2 имеют вид

$$t_1 = -\frac{\mu}{\sigma^2} + \frac{1}{\sigma^2}(\mu^2 - 2\sigma^2\tau)^{1/2}, \qquad t_2 = -\frac{\mu}{\sigma^2} - \frac{1}{\sigma^2}(\mu^2 - 2\sigma^2\tau)^{1/2},$$

где знак квадратного корня определяется так, чтобы действительная часть $\mu^2 - 2\sigma^2\tau$ была положительна.

В предельном случае, когда $B = 0$, A конечно (когда необходимо $\mathbf{M}(z) > 0$, если $\mathbf{M}(n)$ существует), показать, что х. ф. равна A^{-t_1}, а в случае, когда B конечно, $A = 0$ (при $\mathbf{M}(z) < 0$), показать, что х. ф. равна B^{-t_1}.

(Вальд, 1947.)

34.13 В первом из двух предельных случаев предыдущего упражнения доказать, что распределение величины $m = \mu^2 n/2\sigma^2$ дается соотношением

$$dF(m) = \frac{c}{2\Gamma(1/2)\,m^{3/2}} \exp\left(-\frac{c^2}{4m} - m + c\right) dm, \quad 0 \leqslant m \leqslant \infty,$$

где $c = (\mu/\sigma^2)\log A$.

Для больших c доказать, что $2m/c$ приближенно нормально с единичным средним и дисперсией $1/c$.

(Вальд, 1947; там же показано, что, когда A, B конечны, распределение величины n совпадает с распределением взвешенной суммы некоторого числа величин приведенного выше типа.)

34.14 Пусть наблюдаемая величина u имеет экспоненциальное распределение

$$dF = e^{-\lambda u}\lambda\, du, \quad 0 \leqslant u \leqslant \infty.$$

Показать, что последовательный критерий для $\lambda = \lambda_0$ против $\lambda = \lambda_1$ задается неравенствами

$$k_1 + (\lambda_1 - \lambda_0)\sum_{j=1}^{n} u_j \leqslant n \log(\lambda_1/\lambda_0) \leqslant k_2 + (\lambda_1 - \lambda_0)\sum_{j=1}^{n} u_j,$$

где k_1 и k_2 — константы.

Сравнить его с критерием из упражнения 34.3 в предельном случае, когда ω_1 и ω_2 стремятся к нулю так, чтобы $\omega_1 t = \lambda_0$ и $\omega_2 t = \lambda_1$ оставались конечными.

(Энскомб и Пейдж, 1954.)

34.15 Требуется оценить параметр θ с малой дисперсией $a(\theta)/\lambda$, когда λ стремится к бесконечности. Пусть t_m — некоторая несмещенная оценка по выборке фиксированного объема m с дисперсией $v(\theta)/m$; пусть $\gamma_1(t_m) = O(m^{-1/2})$ и $\gamma_2(t_m) = O(m^{-1})$, и пусть $a(t_m)$ и $v(t_m)$ могут быть разложены в ряд, дающий асимптотические средние и стандартные ошибки. Рассмотреть следующее правило двойной выборки.

(а) Берется выборка объема $N\lambda$, и по ней находится t_1, оценка параметра θ.

(б) Берется вторая выборка объема $\max\{0, [\{n_0(t_1) - N\}\lambda]\}$, где $n_0(t_1) = v(t_1)/a(t_1)$. По второй выборке находится t_2, оценка для θ.

(в) Пусть $t = \dfrac{Nt_1 + \{n_0(t_1) - N\}t_2}{n_0(t_1)}$, если $n_0(t_1) \geqslant N$.

(г) Предположим, что $N < n_0(\theta)$ и распределение величины $m_0(t_1) = 1/n_0(t_1)$ таково, что событием $n_0(t_1) < N$ можно пренебречь.

Доказать, что при этом правиле

$$\mathbf{M}(t) = \theta + O(\lambda^{-1}), \quad \mathbf{D}t = a(\theta)\lambda^{-1}\{1 + O(\lambda^{-1})\}.$$

(Д. Кокс, 1952c.)

34.16 Взять ту же процедуру, как и в предыдущем упражнении, заменив лишь $n_0(t_1)$ на

$$n(t_1) = n_0(t_1)\left\{1 + \frac{b(t_1)}{\lambda}\right\}.$$

Показать, что

$$\mathbf{M}(t) = \begin{cases} \theta + m_0'(\theta)v(\theta)\lambda^{-1}, & \text{если } N \leqslant n(t_1), \\ 0 & \text{в остальных случаях.} \end{cases}$$

Положить

$$t' = \begin{cases} t - m_0'(t)v(t)\lambda^{-1}, & \text{если } N \leqslant n(t_1), \\ 0 & \text{в остальных случаях} \end{cases}$$

и показать, что t' имеет смещение $O(\lambda^{-2})$.

Доказать далее, что если мы положим

$$b(\theta) = n_0(\theta)\, v(\theta) \{2m_0(\theta) m_0'(\theta) \gamma_1(\theta) v^{-1/2}(\theta) + m_0'^2(\theta) + 2m_0(\theta) m_0''(\theta) + m_0''(\theta)/(2N)\},$$

то

$$\mathrm{D}t' = a(\theta)\lambda^{-1} + O(\lambda^{-3}).$$

(Д. Кокс, 1952c.)

34.17 Применяя упражнение 34.15 к биномиальному распределению с

$$a(\omega) = a\omega^2, \qquad v(\omega) = \omega(1-\omega), \qquad \gamma_1(\omega) = \frac{-(1-2\omega)}{\{\omega(1-\omega)\}^{1/2}},$$

показать, что общий объем выборки равен

$$n(t_1) = \frac{1-t_1}{at_1} + \frac{3}{t_1(1-t_1)} + \frac{1}{aNt_1}$$

и оценка $t' = t - \dfrac{at}{1-t}$.

Таким образом, N должно быть выбрано возможно большим, но не превосходящим $(1-\omega)/(a\omega)$.

(Д. Кокс, 1952c.)

34.18 Используя пример 34.9, показать, что

$$K(\sigma) = \left\{\left(\frac{1-\beta}{\alpha}\right)^h - 1\right\} \Big/ \left\{\left(\frac{1-\beta}{\alpha}\right)^h - \left(\frac{\beta}{1-\alpha}\right)^h\right\},$$

где h определяется из уравнения

$$\sigma\left(\frac{\sigma_1}{\sigma_0}\right)^h = \left\{\frac{h}{\sigma_1^2} - \frac{h}{\sigma_0^2} + \frac{1}{\sigma^2}\right\}^{-1/2}$$

при условии, что выражение в скобках справа положительно. Как, пользуясь этим результатом, строить кривую OX?

(Вальд, 1947.)

34.19 В условиях предыдущего упражнения вывести выражение для СОВ

$$\frac{K(\sigma)\{h_0 - h_1\} + h_1}{\sigma^2 - \gamma}, \qquad \text{где} \qquad \gamma = \log\left(\frac{\sigma_1^2}{\sigma_0^2}\right) \Big/ \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right).$$

(Вальд, 1947.)

34.20 Обосновать утверждение последнего предложения примера 34.9, дающее критерий для дисперсий нормального распределения, когда среднее неизвестно.

(Гиршик, 1946.)

34.21 Заменим в **34.10** $f(x, \theta)$ на $f(x, \theta, \varphi)$, где φ — мешающий параметр, так, чтобы гипотезы H_0 и H_1 стали сложными. Пусть $(\hat{\theta}_n, \hat{\varphi}_n)$ — МП-оценка для (θ, φ) по n наблюдениям. Показать, что если θ_0 и θ_1 отличаются от истинного значения θ на величину порядка $n^{-1/2}$, то ПКОВ, основанный на

$$t_n(\hat{\varphi}) = \log L_n(x, \theta_1, \hat{\varphi}) - \log L_n(x, \theta_0, \hat{\varphi}),$$

асимптотически эквивалентен (34.18) для простых гипотез, когда φ известно, тогда и только тогда, когда

$$\frac{1}{n}\frac{\partial^2 \log L_n(x, \theta, \varphi)}{\partial\theta\,\partial\varphi} \to 0$$

по вероятности, т. е. если $\hat{\theta}$ и $\hat{\varphi}$ асимптотически независимы. В любом случае показать, что

$$t_n(\hat{\varphi}) \sim (\theta_1 - \theta_0)\left\{\hat{\theta} - \frac{1}{2}(\theta_0 + \theta_1)\right\}\mathbf{M}\left\{-\frac{\partial^2 \log L_n}{\partial\theta^2}\right\},$$

так что $T_n = n\left\{\hat{\theta} - \frac{1}{2}(\theta_0 + \theta_1)\right\}$ может использоваться в качестве статистики критерия со средним и дисперсией, сразу получающимися из таковых для $\hat{\theta}$.

(Д. Кокс, 1963.)

34.22 В случае выборки из распределения $f(x|\theta) = g(x)/h(\theta)$, $a \leqslant x \leqslant \theta$, доказать, что ПКОВ для гипотезы H_0: $\theta = \theta_0$ против H_1: $\theta = \theta_1$ $(0 < \theta_0 < \theta_1)$ имеет $\alpha = 0$ и задается правилом: принять H_1, если $x_n > \theta_0$, $n = 1, 2, \ldots$; принять H_0, если $x_0 \leqslant \theta_0$ и $(\theta_0/\theta_1)^n \leqslant \beta$; в других случаях продолжать выбор.

Непосредственно показать, что, когда выполняется H_0, объем выборки равен постоянной $n_0 = \log\beta/\log(\theta_0/\theta_1)$, а когда выполняется H_1, СОВ в точности равен $(1-\beta)\Big/\left(1 - \frac{\theta_0}{\theta_1}\right)$, если n_0 целое. Проверить эти формулы для СОВ из (34.43).

ПРИЛОЖЕНИЕ

Таблица 1

Функция плотности нормального распределения

$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ с первыми и вторыми разностями

x	y	Δ^1 (−)	Δ^2	x	y	Δ^1 (−)	Δ^2
0,0	0,39894	199	−392	2,5	0,01753	395	+79
0,1	0,39695	591	−374	2,6	0,01358	316	+66
0,2	0,39104	965	−347	2,7	0,01042	250	+53
0,3	0,38139	1312	−303	2,8	0,00792	197	+45
0,4	0,36827	1620	−265	2,9	0,00595	152	+36
0,5	0,35207	1885	−212	3,0	0,00443	116	+27
0,6	0,33322	2097	−159	3,1	0,00327	89	+23
0,7	0,31225	2256	−101	3,2	0,00238	66	+17
0,8	0,28969	2360	−52	3,3	0,00172	49	+13
0,9	0,26609	2412	−0	3,4	0,00123	36	+10
1,0	0,24197	2412	+46	3,5	0,00087	26	+7
1,1	0,21785	2366	+84	3,6	0,00061	19	+6
1,2	0,19419	2282	+118	3,7	0,00042	13	+4
1,3	0,17137	2164	+143	3,8	0,00029	9	+2
1,4	0,14973	2021	+161	3,9	0,00020	7	+3
1,5	0,12952	1860	+173	4,0	0,00013	4	—
1,6	0,11092	1687	+177	4,1	0,00009	3	—
1,7	0,09405	1510	+177	4,2	0,00006	2	—
1,8	0,07895	1333	+170	4,3	0,00004	2	—
1,9	0,06562	1163	+162	4,4	0,00002	—	—
2,0	0,05399	1001	+150	4,5	0,00002	—	—
2,1	0,04398	851	+137	4,6	0,00001	—	—
2,2	0,03547	714	+120	4,7	0,00001	—	—
2,3	0,02833	594	+108	4,8	0,00000	—	—
2,4	0,02239	486	+91				

Таблица 2

Функция нормального распределения

В таблице даны значения $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}x^2} dx$, т. е. значения площади под кривой $y = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, лежащей левее точки x; например, для $x = 1{,}86\ (= 1{,}5 + 0{,}36)$ $F(1{,}86) = 0{,}9686$

x	0,0+	0,5+	1,0+	1,5+	2,0+	2,5+	3,0+	3,5+
0,00	5000	6915	8413	9332	9772	9^2379	9^2865	9^377
0,01	5040	6950	8438	9345	9778	9^2396	9^2869	9^378

Продолжение

x	0,0+	0,5+	1,0+	1,5+	2,0+	2,5+	3,0+	3,5+
0,02	5080	6985	8461	9357	9783	$9^{2}413$	$9^{2}874$	$9^{3}78$
0,03	5120	7019	8485	9370	9788	$9^{2}430$	$9^{2}878$	$9^{3}79$
0,04	5160	7054	8508	9382	9793	$9^{2}446$	$9^{2}882$	$9^{3}80$
0,05	5199	7088	8531	9394	9798	$9^{2}461$	$9^{2}886$	$9^{3}81$
0,06	5239	7123	8554	9406	9803	$9^{2}477$	$9^{2}889$	$9^{3}81$
0,07	5279	7157	8577	9418	9808	$9^{2}492$	$9^{2}893$	$9^{3}82$
0,08	5319	7190	8599	9429	9812	$9^{2}506$	$9^{2}897$	$9^{3}83$
0,09	5359	7224	8621	9441	9817	$9^{2}520$	$9^{2}900$	$9^{3}83$
0,10	5398	7257	8643	9452	9821	$9^{2}534$	$9^{3}03$	$9^{3}84$
0,11	5438	7291	8665	9463	9826	$9^{2}547$	$9^{3}06$	$9^{3}85$
0,12	5478	7324	8686	9474	9830	$9^{2}560$	$9^{3}10$	$9^{3}85$
0,13	5517	7357	8708	9484	9834	$9^{2}573$	$9^{3}13$	$9^{3}86$
0,14	5557	7389	8729	9495	9838	$9^{2}585$	$9^{3}16$	$9^{3}86$
0,15	5596	7422	8749	9505	9842	$9^{2}598$	$9^{3}18$	$9^{3}87$
0,16	5636	7454	8770	9515	9846	$9^{2}609$	$9^{3}21$	$9^{3}87$
0,17	5675	7486	8790	9525	9850	$9^{2}621$	$9^{3}24$	$9^{3}88$
0,18	5714	7517	8810	9535	9854	$9^{2}632$	$9^{3}26$	$9^{3}88$
0,19	5753	7549	8830	9545	9857	$9^{2}643$	$9^{3}29$	$9^{3}89$
0,20	5793	7580	8849	9554	9861	$9^{2}653$	$9^{3}31$	$9^{3}89$
0,21	5832	7611	8869	9564	9864	$9^{2}664$	$9^{3}34$	$9^{3}90$
0,22	5871	7642	8888	9573	9868	$9^{2}674$	$9^{3}36$	$9^{3}90$
0,23	5910	7673	8907	9582	9871	$9^{2}683$	$9^{3}38$	$9^{1}04$
0,24	5948	7704	8925	9591	9875	$9^{2}693$	$9^{3}40$	$9^{1}08$
0,25	5987	7738	8944	9599	9878	$9^{2}702$	$9^{3}42$	$9^{1}12$
0,26	6026	7764	8962	9608	9881	$9^{2}711$	$9^{3}44$	$9^{1}15$
0,27	6064	7794	8980	9616	9884	$9^{2}720$	$9^{3}46$	$9^{1}18$
0,28	6103	7823	8997	9625	9887	$9^{2}728$	$9^{3}48$	$9^{1}22$
0,29	6141	7852	9015	9633	9890	$9^{2}736$	$9^{3}50$	$9^{1}25$
0,30	6179	7881	9032	9641	9893	$9^{2}744$	$9^{3}52$	$9^{1}28$
0,31	6217	7910	9049	9649	9896	$9^{2}752$	$9^{3}53$	$9^{1}31$
0,32	6255	7939	9066	9656	9898	$9^{2}760$	$9^{3}55$	$9^{1}33$
0,33	6293	7967	9082	9664	9901	$9^{2}767$	$9^{3}57$	$9^{1}36$
0,34	6331	7995	9099	9671	9904	$9^{2}774$	$9^{3}58$	$9^{1}39$
0,35	6368	8023	9115	9678	9906	$9^{2}781$	$9^{3}60$	$9^{1}41$
0,36	6406	8051	9131	9686	9909	$9^{2}788$	$9^{3}61$	$9^{1}43$
0,37	6443	8078	9147	9693	9911	$9^{2}795$	$9^{3}62$	$9^{1}46$
0,38	6480	8106	9162	9699	9913	$9^{2}801$	$9^{3}64$	$9^{1}48$
0,39	6517	8133	9177	9706	9916	$9^{2}807$	$9^{3}65$	$9^{1}50$
0,40	6554	8159	9192	9713	9918	$9^{2}813$	$9^{3}66$	$9^{1}52$
0,41	6591	8186	9207	9719	9920	$9^{2}819$	$9^{3}68$	$9^{1}54$
0,42	6628	8212	9222	9726	9922	$9^{2}825$	$9^{3}69$	$9^{1}56$
0,43	6664	8238	9236	9732	9925	$9^{2}831$	$9^{3}70$	$9^{1}58$
0,44	6700	8264	9251	9738	9927	$9^{2}836$	$9^{3}71$	$9^{1}59$
0,45	6736	8289	9265	9744	9929	$9^{2}841$	$9^{3}72$	$9^{1}61$
0,46	6772	8315	9279	9750	9931	$9^{2}846$	$9^{3}73$	$9^{1}63$
0,47	6808	8340	9292	9756	9932	$9^{2}851$	$9^{3}74$	$9^{1}64$
0,48	6844	8365	9306	9761	9934	$9^{2}856$	$9^{3}75$	$9^{1}66$
0,49	6879	8389	9319	9767	9936	$9^{2}861$	$9^{3}76$	$9^{1}67$

Примечание. Знаки целой части (т. е. 0) в таблице опущены. Повторяющиеся девятки заменены степенями; например, $9^{3}71$ означает 0,99971.

Квантили χ^2-распределения

Таблица 3

(Заимствованы из таблицы III книги сэра Рональда Фишера *Statistical Methods for Research Workers*, Oliver and Boyd Ltd., Edinburgh.)

$P=1-F$	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01
$\nu=$ 1	$0,0^3157$	$0,0^3628$	$0,0^2393$	0,0158	0,0642	0,148	0,455	1,074	1,642	2,706	3,841	5,412	6,635
2	0,0201	0,0404	0,103	0,211	0,446	0,713	1,386	2,408	3,219	4,605	5,991	7,824	9,210
3	0,115	0,185	0,352	0,584	1,005	1,424	2,366	3,665	4,642	6,251	7,815	9,837	11,345
4	0,297	0,429	0,711	1,064	1,649	2,195	3,357	4,878	5,989	7,779	9,488	11,668	13,277
5	0,554	0,752	1,145	1,160	2,343	3,000	4,351	6,064	7,289	9,236	11,070	13,388	15,086
6	0,872	1,134	1,635	2,204	3,070	3,828	5,348	7,231	8,558	10,645	12,592	15,033	16,812
7	1,239	1,564	2,167	2,833	3,822	4,671	6,346	8,383	9,803	12,017	14,067	16,622	18,475
8	1,646	2,032	2,733	3,490	4,594	5,527	7,344	9,524	11,030	13,362	15,507	18,168	20,090
9	2,088	2,532	3,325	4,168	5,380	6,393	8,343	10,656	12,242	14,684	16,919	19,679	21,666
10	2,358	3,059	3,940	4,865	6,179	7,267	9,342	11,781	13,442	15,987	18,307	21,161	23,209
11	3,053	3,609	4,575	5,578	6,989	8,148	10,341	12,899	14,631	17,275	19,675	22,618	24,725
12	3,571	4,178	5,226	6,304	7,807	9,034	11,340	14,011	15,821	18,549	21,026	24,054	26,217
13	4,107	4,765	5,892	7,042	8,634	9,926	12,340	15,119	16,985	19,812	22,362	25,472	27,688
14	4,660	5,368	6,571	7,790	9,467	10,821	13,339	16,222	18,151	21,064	23,685	26,873	29,141
15	5,229	5,985	7,261	8,547	10,307	11,721	14,339	17,322	19,311	22,307	24,996	28,259	30,578
16	5,812	6,614	7,962	9,312	11,152	12,624	15,338	18,418	20,465	23,542	26,296	29,633	32,000
17	6,408	7,255	8,672	10,085	12,002	13,531	16,338	19,511	21,615	24,769	27,587	30,995	33,409
18	7,015	7,906	9,390	10,865	12,857	14,440	17,338	20,601	22,760	25,989	28,869	32,346	34,805
19	7,633	8,567	10,117	11,651	13,716	15,352	18,338	21,689	23,900	27,204	30,144	33,687	36,191
20	8,260	9,237	10,851	12,443	14,578	16,266	19,337	22,775	25,038	28,412	31,410	35,020	37,566
21	8,897	9,915	11,591	13,240	15,445	17,182	20,337	23,858	26,171	29,615	32,671	36,343	38,932
22	9,542	10,600	12,338	14,041	16,314	18,101	21,337	24,939	27,301	30,813	33,924	37,659	40,289
23	10,196	11,293	13,091	14,848	17,187	19,021	22,337	26,018	28,429	32,007	35,172	38,968	41,638
24	10,856	11,992	13,848	15,659	18,062	19,943	23,337	27,096	29,553	33,196	36,415	40,270	42,980
25	11,524	12,697	14,611	16,473	18,940	20,867	24,337	28,172	30,675	34,382	37,652	41,566	44,314
26	12,198	13,409	15,379	17,292	19,820	21,792	25,336	29,246	31,795	35,563	38,885	42,856	45,642
27	12,879	14,125	16,151	18,114	20,703	22,719	26,336	30,319	32,912	36,741	40,113	44,140	46,963
28	13,565	14,847	16,928	18,939	21,588	23,647	27,336	31,391	34,027	37,916	41,337	45,419	48,278
29	14,256	15,574	17,708	19,768	22,475	24,577	28,336	32,461	35,139	39,087	42,557	46,693	49,588
30	14,953	16,306	18,493	20,599	23,364	25,508	29,336	33,530	36,250	40,256	43,773	47,962	50,892

Примечание. Для значений ν, бо́льших 30, величину $\sqrt{2\chi^2}$ можно считать распределенной нормально со средним $\sqrt{2\nu-1}$ и единичной дисперсией.

Таблица 4

Функция χ^2-распределения с одной степенью свободы
для $0 \leqslant \chi^2 \leqslant 1$ с шагом 0,01

χ^2	P	Δ	χ^2	P	Δ	χ^2	P	Δ
0	1,00000	7966	0,34	0,55983	572	0,68	0,40959	343
0,01	0,92034	3280	0,35	0,55411	560	0,69	0,40616	338
0,02	0,88754	2505	0,36	0,54851	551	0,70	0,40278	334
0,03	0,86249	2101	0,37	0,54300	540	0,71	0,39944	330
0,04	0,84148	1842	0,38	0,53760	530	0,72	0,39614	326
0,05	0,82306	1656	0,39	0,53230	521	0,73	0,39288	322
0,06	0,80650	1516	0,40	0,52709	512	0,74	0,38966	318
0,07	0,79134	1404	0,41	0,52197	503	0,75	0,38648	315
0,08	0,77730	1312	0,42	0,51694	495	0,76	0,38333	311
0,09	0,76418	1235	0,43	0,51199	487	0,77	0,38022	308
0,10	0,75183	1169	0,44	0,50712	479	0,78	0,37714	304
0,11	0,74014	1111	0,45	0,50233	471	0,79	0,37410	301
0,12	0,72903	1060	0,46	0,49762	463	0,80	0,37109	297
0,13	0,71843	1015	0,47	0,49299	457	0,81	0,36812	294
0,14	0,70828	974	0,48	0,48842	449	0,82	0,36518	291
0,15	0,69854	938	0,49	0,48393	443	0,83	0,36227	287
0,16	0,68916	905	0,50	0,47950	436	0,84	0,35940	285
0,17	0,68011	874	0,51	0,47514	430	0,85	0,35655	281
0,18	0,67137	845	0,52	0,47084	423	0,86	0,35374	278
0,19	0,66292	820	0,53	0,46661	418	0,87	0,35096	276
0,20	0,65472	795	0,54	0,46243	411	0,88	0,34820	272
0,21	0,64677	773	0,55	0,45832	406	0,89	0,34548	270
0,22	0,63904	752	0,56	0,45426	400	0,90	0,34278	267
0,23	0,63152	731	0,57	0,45026	395	0,91	0,34011	264
0,24	0,62421	713	0,58	0,44631	389	0,92	0,33747	261
0,25	0,61708	696	0,59	0,44242	384	0,93	0,33486	258
0,26	0,61012	679	0,60	0,43858	379	0,94	0,33228	256
0,27	0,60333	663	0,61	0,43479	374	0,95	0,32972	253
0,28	0,59670	648	0,62	0,43105	369	0,96	0,32719	251
0,29	0,59022	634	0,63	0,42736	365	0,97	0,32468	248
0,30	0,58388	620	0,64	0,42371	360	0,98	0,32220	246
0,31	0,57768	607	0,65	0,42011	355	0,99	0,31974	243
0,32	0,57161	595	0,66	0,41656	351	1,00	0,31731	241
0,33	0,56566	583	0,67	0,41305	346			

Таблица 4б

Функция χ^2-распределения с одной степенью свободы
для $1 \leqslant \chi^2 \leqslant 10$ с шагом 0,1

χ^2	P	Δ	χ^2	P	Δ	χ^2	P	Δ
1,0	0,31731	2304	1,4	0,23672	1605	1,8	0,17971	1163
1,1	0,29427	2095	1,5	0,22067	1477	1,9	0,16808	1078
1,2	0,27332	1911	1,6	0,20590	1361	2,0	0,15730	1000
1,3	0,25421	1749	1,7	0,19229	1258	2,1	0,14730	929

Продолжение

χ^2	P	Δ	χ^2	P	Δ	χ^2	P	Δ
2,2	0,13801	864	4,9	0,02686	151	7,6	0,00584	32
2,3	0,12937	803	5,0	0,02535	142	7,7	0,00552	30
2,4	0,12134	749	5,1	0,02393	134	7,8	0,00522	28
2,5	0,11385	699	5,2	0,02259	126	7,9	0,00494	26
2,6	0,10686	651	5,3	0,02133	119	8,0	0,00468	25
2,7	0,10035	609	5,4	0,02014	112	8,1	0,00443	24
2,8	0,09426	568	5,5	0,01902	106	8,2	0,00419	23
2,9	0,08858	532	5,6	0,01796	99	8,3	0,00396	21
3,0	0,08326	497	5,7	0,01697	94	8,4	0,00375	20
3,1	0,07829	465	5,8	0,01603	89	8,5	0,00355	19
3,2	0,07364	436	5,9	0,01514	83	8,6	0,00336	18
3,3	0,06928	408	6,0	0,01431	79	8,7	0,00318	17
3,4	0,06520	383	6,1	0,01352	74	8,8	0,00301	16
3,5	0,06137	359	6,2	0,01278	71	8,9	0,00285	15
3,6	0,05778	337	6,3	0,01207	66	9,0	0,00270	14
3,7	0,05441	316	6,4	0,01141	62	9,1	0,00256	14
3,8	0,05125	296	6,5	0,01079	59	9,2	0,00242	13
3,9	0,04829	279	6,6	0,01020	56	9,3	0,00229	12
4,0	0,04550	262	6,7	0,00964	52	9,4	0,00217	12
4,1	0,04288	246	6,8	0,00912	50	9,5	0,00205	10
4,2	0,04042	231	6,9	0,00862	47	9,6	0,00195	11
4,3	0,03811	217	7,0	0,00815	44	9,7	0,00184	10
4,4	0,03594	205	7,1	0,00771	42	9,8	0,00174	9
4,5	0,03389	192	7,2	0,00729	39	9,9	0,00165	8
4,6	0,03197	181	7,3	0,00690	38	10,0	0,00157	8
4,7	0,03016	170	7,4	0,00652	35			
4,8	0,02846	160	7,5	0,00617	33			

Таблица 5

Квантили t-распределения

(Заимствованы из таблиц сэра Рональда Фишера и Ф. Иэйтса: *Statistical Tables for Biological, Medical and Agricultural Research*, Oliver and Boyd Ltd., Edinburgh.)

$P=2(1-F)$	0,9	0,8	0,7	0,6	0,5	0,4	0,3
$\nu=$ 1	0,158	0,325	0,510	0,727	1,000	1,376	1,963
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093

Продолжение

$P=2(1-F)$	0,9	0,8	0,7	0,6	0,5	0,4	0,3
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050
60	0,126	0,254	0,387	0,527	0,679	0,848	1,046
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036

$P=2(1-F)$	0,2	0,1	0,05	0,02	0,01	0,001
$\nu=$ 1	3,078	6,314	12,706	31,821	63,657	636,619
2	1,886	2,920	4,303	6,965	9,925	31,598
3	1,638	2,353	3,182	4,541	5,841	12,924
4	1,533	2,132	2,776	3,747	4,604	8,610
5	1,476	2,015	2,571	3,365	4,032	6,869
6	1,440	1,943	2,447	3,143	3,707	5,959
7	1,415	1,895	2,365	2,998	3,499	5,408
8	1,397	1,860	2,306	2,896	3,355	5,041
9	1,383	1,833	2,262	2,821	3,250	4,781
10	1,372	1,812	2,228	2,764	3,169	4,587
11	1,363	1,796	2,201	2,718	3,106	4,437
12	1,356	1,782	2,179	2,681	3,055	4,318
13	1,350	1,771	2,160	2,650	3,012	4,221
14	1,345	1,761	2,145	2,624	2,977	4,140
15	1,341	1,753	2,131	2,602	2,947	4,073

Продолжение

$P=2(1-F)$	0,2	0,1	0,05	0,02	0,01	0,001
16	1,337	1,746	2,120	2,583	2,921	4,015
17	1,333	1,740	2,110	2,567	2,898	3,965
18	1,330	1,734	2,101	2,552	2,878	3,922
19	1,328	1,729	2,093	2,539	2,861	3,883
20	1,325	1,725	2,086	2,528	2,845	3,850
21	1,323	1,721	2,080	2,518	2,831	3,819
22	1,321	1,717	2,074	2,508	2,819	3,792
23	1,319	1,714	2,069	2,500	2,807	3,767
24	1,318	1,711	2,064	2,492	2,797	3,745
25	1,316	1,708	2,060	2,485	2,787	3,725
26	1,315	1,706	2,056	2,479	2,779	3,707
27	1,314	1,703	2,052	2,473	2,771	3,690
28	1,313	1,701	2,048	2,467	2,763	3,674
29	1,311	1,699	2,045	2,462	2,756	3,659
30	1,310	1,697	2,042	2,457	2,750	3,646
40	1,303	1,684	2,021	2,423	2,704	3,551
60	1,296	1,671	2,000	2,390	2,660	3,460
120	1,289	1,658	1,980	2,358	2,617	3,373
∞	1,282	1,645	1,960	2,326	2,576	3,291

Таблица 6

5%-ные точки *z*-распределения (точки, в которых ф. р. = 0,95)

(Заимствованы из таблицы VI книги сэра Рональда Фишера *Statistical Methods for Research Workers*, Oliver and Boyd Ltd., Edinburgh.)

Значения ν_2	Значения ν_1									
	1	2	3	4	5	6	8	12	24	∞
1	2,5421	2,6479	2,6870	2,7071	2,7194	2,7276	2,7380	2,7484	2,7588	2,7693
2	1,4592	1,4722	1,4765	1,4787	1,4800	1,4808	1,4819	1,4830	1,4840	1,4851
3	1,1577	1,1284	1,1137	1,1051	1,0994	1,0953	1,0899	1,0842	1,0781	1,0716
4	1,0212	0,9690	0,9429	0,9272	0,9168	0,9093	0,8993	0,8885	0,8767	0,8639
5	0,9441	0,8777	0,8441	0,8236	0,8097	0,7997	0,7862	0,7714	0,7550	0,7368
6	0,8948	0,8188	0,7798	0,7558	0,7394	0,7274	0,7112	0,6931	0,6729	0,6499
7	0,8606	0,7777	0,7347	0,7080	0,6896	0,6761	0,6576	0,6369	0,6134	0,5862
8	0,8355	0,7475	0,7014	0,6725	0,6525	0,6378	0,6175	0,5945	0,5682	0,5371
9	0,8163	0,7242	0,6757	0,6450	0,6238	0,6080	0,5862	0,5613	0,5324	0,4979
10	0,8012	0,7058	0,6553	0,6232	0,6009	0,5843	0,5611	0,5346	0,5035	0,4657
11	0,7889	0,6909	0,6387	0,6055	0,5822	0,5648	0,5406	0,5126	0,4795	0,4387
12	0,7788	0,6786	0,6250	0,5907	0,5666	0,5487	0,5234	0,4941	0,4592	0,4156
13	0,7703	0,6682	0,6134	0,5783	0,5535	0,5350	0,5089	0,4785	0,4419	0,3957
14	0,7630	0,6594	0,6036	0,5677	0,5423	0,5233	0,4964	0,4649	0,4269	0,3782
15	0,7568	0,6518	0,5950	0,5585	0,5326	0,5131	0,4855	0,4532	0,4138	0,3628

Продолжение

Значения ν_2	Значения ν_1									
	1	2	3	4	5	6	8	12	24	∞
16	0,7514	0,6451	0,5876	0,5505	0,5241	0,5042	0,4760	0,4428	0,4022	0,3490
17	0,7466	0,6393	0,5811	0,5434	0,5166	0,4964	0,4676	0,4337	0,3919	0,3366
18	0,7424	0,6341	0,5753	0,5371	0,5099	0,4894	0,4602	0,4255	0,3827	0,3253
19	0,7386	0,6295	0,5701	0,5315	0,5040	0,4832	0,4535	0,4182	0,3743	0,3151
20	0,7352	0,6254	0,5654	0,5265	0,4986	0,4776	0,4474	0,4116	0,3668	0,3057
21	0,7322	0,6216	0,5612	0,5219	0,4938	0,4725	0,4420	0,4055	0,3599	0,2971
22	0,7294	0,6182	0,5574	0,5178	0,4894	0,4679	0,4370	0,4001	0,3536	0,2892
23	0,7269	0,6151	0,5540	0,5140	0,4854	0,4636	0,4325	0,3950	0,3478	0,2818
24	0,7246	0,6123	0,5508	0,5106	0,4817	0,4598	0,4283	0,3904	0,3425	0,2749
25	0,7225	0,6097	0,5478	0,5074	0,4783	0,4562	0,4244	0,3862	0,3376	0,2685
26	0,7205	0,6073	0,5451	0,5045	0,4752	0,4529	0,4209	0,3823	0,3330	0,2625
27	0,7187	0,6051	0,5427	0,5017	0,4723	0,4499	0,4176	0,3786	0,3287	0,2569
28	0,7171	0,6030	0,5403	0,4992	0,4696	0,4471	0,4146	0,3752	0,3248	0,2516
29	0,7155	0,6011	0,5382	0,4969	0,4671	0,4444	0,4117	0,3720	0,3211	0,2466
30	0,7141	0,5994	0,5362	0,4947	0,4648	0,4420	0,4090	0,3691	0,3176	0,2419
60	0,6933	0,5738	0,5073	0,4632	0,4311	0,4064	0,3702	0,3255	0,2654	0,1644
∞	0,6729	0,5486	0,4787	0,4319	0,3974	0,3706	0,3309	0,2804	0,2085	0

Таблица 7

5%-ные точки *F*-распределения (точки, в которых ф. р. = 0,95)

(Заимствованы из таблиц сэра Рональда Фишера и Ф. Иэйтса: *Statistical Tables for Biological, Medical and Agricultural Research*, Oliver and Boyd Ltd., Edinburgh.)

ν_2 \ ν_1	1	2	3	4	5	6	8	12	24	∞
1	161,40	199,50	215,70	224,60	230,20	234,00	238,90	243,90	249,00	254,30
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07

Продолжение

ν_2 \ ν_1	1	2	3	4	5	6	8	12	24	∞
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,02	1,83	1,61	1,25
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00

Примечание. Нижняя 5%-ная точка представляет собой обратную величину того табличного значения, которое получится, если ν_1 и ν_2 поменять местами, т. е. в качестве ν_1 всегда следует выбирать количество степеней свободы, соответствующее наибольшей выборочной дисперсии.

Таблица 8

1%-ные точки *z*-распределения (точки, в которых ф. р. = 0,99)

(Заимствованы из таблицы VI книги сэра Рональда Фишера *Statistical Methods for Research Workers*, Oliver and Boyd Ltd., Edinburgh.)

Значения ν_2 \ Значения ν_1	1	2	3	4	5	6	8	12	24	∞
1	4,1535	4,2585	4,2974	4,3175	4,3297	4,3379	4,3482	4,3585	4,3689	4,3794
2	2.2950	2,2976	2,2984	2,2988	2,2991	2,2992	2,2994	2,2997	2,2999	2,3001
3	1,7649	1,7140	1,6915	1,6786	1,6703	1,6645	1,6569	1,6489	1,6404	1,6314
4	1,5270	1,4452	1,4075	1,3856	1,3711	1,3609	1,3473	1,3327	1,3170	1,3000
5	1,3943	1,2929	1,2449	1,2164	1,1974	1,1838	1,1656	1,1457	1,1239	1,0997
6	1,3103	1,1955	1,1401	1,1068	1,0843	1,0680	1,0460	1,0218	0,9948	0,9643
7	1,2526	1,1281	1,0672	1,0300	1,0048	0,9864	0,9614	0,9335	0,9020	0,8658
8	1,2103	1,0787	1,0135	0,9734	0,9459	0,9259	0,8983	0,8673	0,8319	0,7904
9	1,1786	1,0411	0,9724	0,9299	0,9006	0,8791	0,8494	0,8157	0,7769	0,7305
10	1,1535	1,0114	0,9399	0,8954	0,8646	0,8419	0,8104	0,7744	0,7324	0,6816

Продолжение

Значения ν_2	Значения ν_1: 1	2	3	4	5	6	8	12	24	∞
11	1,1333	0,9874	0,9136	0,8674	0,8354	0,8116	0,7785	0,7405	0,6958	0,6408
12	1,1166	0,9677	0,8919	0,8443	0,8111	0,7864	0,7520	0,7122	0,6649	0,6061
13	1,1027	0,9511	0,8737	0,8248	0,7907	0,7652	0,7295	0,6882	0,6386	0,5761
14	1,0909	0,9370	0,8581	0,8082	0,7732	0,7471	0,7103	0,6675	0,6159	0,5500
15	1,0807	0,9249	0,8448	0,7939	0,7582	0,7314	0,6937	0,6496	0,5961	0,5269
16	1,0719	0,9144	0,8331	0,7814	0,7450	0,7177	0,6791	0,6339	0,5786	0,5064
17	1,0641	0,9051	0,8229	0,7705	0,7335	0,7057	0,6663	0,6199	0,5630	0,4879
18	1,0572	0,8970	0,8138	0,7607	0,7232	0,6950	0,6549	0,6075	0,5491	0,4712
19	1,0511	0,8897	0,8057	0,7521	0,7140	0,6854	0,6447	0,5964	0,5366	0,4560
20	1,0457	0,8831	0,7985	0,7443	0,7058	0,6768	0,6355	0,5864	0,5253	0,4421
21	1,0408	0,8772	0,7920	0,7372	0,6984	0,6690	0,6272	0,5773	0,5150	0,4294
22	1,0363	0,8719	0,7860	0,7309	0,6916	0,6620	0,6196	0,5691	0,5056	0,4176
23	1,0322	0,8670	0,7806	0,7251	0,6855	0,6555	0,6127	0,5615	0,4969	0,4068
24	1,0285	0,8626	0,7757	0,7197	0,6799	0,6496	0,6064	0,5545	0,4890	0,3967
25	1,0251	0,8585	0,7712	0,7148	0,6747	0,6442	0,6006	0,5481	0,4816	0,3872
26	1,0220	0,8548	0,7670	0,7103	0,6699	0,6392	0,5952	0,5422	0,4748	0,3784
27	1,0191	0,8513	0,7631	0,7062	0,6655	0,6346	0,5902	0,5367	0,4685	0,3701
28	1,0164	0,8481	0,7595	0,7023	0,6614	0,6303	0,5856	0,5316	0,4626	0,3624
29	1,0139	0,8451	0,7562	0,6987	0,6576	0,6263	0,5813	0,5269	0,4570	0,3550
30	1,0116	0,8423	0,7531	0,6954	0,6540	0,6226	0,5773	0,5224	0,4519	0,3481
60	0,9784	0,8025	0,7086	0,6472	0,6028	0,5687	0,5189	0,4574	0,3746	0,2352
∞	0,9462	0,7636	0,6651	0,5999	0,5522	0,5152	0,4604	0,3908	0,2913	0

Таблица 9

1%-ные точки *F*-распределения (точки, в которых ф. р. = 0.99)

(Заимствованы из книги сэра Рональда Фишера и Ф. Иэйтса: *Statistical Tables for Biological, Medical and Agricultural Research*, Oliver and Boyd Ltd., Edinburgh.)

ν_2 \ ν_1	1	2	3	4	5	6	8	12	24	∞
1	4052	4999	5403	5625	5764	5859	5981	6106	6234	6366
2	98,49	99,00	99,17	99,25	99,30	99,33	99,36	99,42	99,46	99,50
3	34,12	30,81	29,46	28,71	28,24	27,91	27,49	27,05	26,60	26,12
4	21,20	18,00	16,69	15,98	15,52	15,21	14,80	14,37	13,93	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,27	9,89	9,47	9,02
6	13,74	10,92	9,78	9,15	8,75	8,47	8,10	7,72	7,31	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,84	6,47	6,07	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,03	5,67	5,28	4,86
9	10,56	8,02	6,99	6,42	6,06	5,80	5,47	5,11	4,73	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,06	4,71	4,33	3,91

Продолжение

ν_2 \ ν_1	1	2	3	4	5	6	8	12	24	∞
11	9,65	7,20	6,22	5,67	5,32	5,07	4,74	4,40	4,02	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,50	4,16	3,78	3,36
13	9,07	6,70	5,74	5,20	4,86	4,62	4,30	3,96	3,59	3,16
14	8,86	6,51	5,56	5,03	4,69	4,46	4,14	3,80	3,43	3,00
15	8,68	6,36	5,42	4,89	4,56	4,32	4,00	3,67	3,29	2,87
16	8,53	6,23	5,29	4,77	4,44	4,20	3,89	3,55	3,18	2,75
17	8,40	6,11	5,18	4,67	4,34	4,10	3,79	3,45	3,08	2,65
18	8,28	6,01	5,09	4,58	4,25	4,01	3,71	3,37	3,00	2,57
19	8,18	5,93	5,01	4,50	4,17	3,94	3,63	3,30	2,92	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,56	3,23	2,86	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,51	3,17	2,80	2,36
22	7,94	5,72	4,82	4,31	3,99	3,76	3,45	3,12	2,75	2,31
23	7,88	5,66	4,76	4,26	3,94	3,71	3,41	3,07	2,70	2,26
24	7,82	5,61	4,72	4,22	3,90	3,67	3,36	3,03	2,66	2,21
25	7,77	5,57	4,68	4,18	3,86	3,63	3,32	2,99	2,62	2,17
26	7,72	5,53	4,64	4,14	3,82	3,59	3,29	2,96	2,58	2,13
27	7,68	5,49	4,60	4,11	3,78	3,56	3,26	2,93	2,55	2,10
28	7,64	5,45	4,57	4,07	3,75	3,53	3,23	2,90	2,52	2,06
29	7,60	5,42	4,54	4,04	3,73	3,50	3,20	2,87	2,49	2,03
30	7,56	5,39	4,51	4,02	3,70	3,47	3,17	2,84	2,47	2,01
40	7,31	5,18	4,31	3,83	3,51	3,29	2.99	2,66	2,29	1,80
60	7,08	4,98	4,13	3,65	3,34	3,12	2,82	2,50	2,12	1,60
120	6,85	4,79	3,95	3,48	3,17	2,96	2,66	2,34	1,95	1,38
∞	6,64	4,60	3,78	3,32	3,02	2,80	2,51	2,18	1,79	1,00

Примечание. Нижняя 1%-ная точка представляет собой обратную величину того табличного значения, которое получится, если ν_1 и ν_2 поменять местами, т. е. в качестве ν_1 всегда следует выбирать количество степеней свободы, соответствующее наибольшей выборочной дисперсии.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

Аллан (Allan F. E.) (1930), The general form of the orthogonal polynomials for simple series with proofs of their simple properties, *Proc. Roy. Soc. Edin.* **A, 50,** 310.

Аллен (Allen H. V.) (1938), A theorem concerning the linearity of regression, *Statist. Res. Mem.*, **2,** 60.

Андерсон (Anderson T. W.) (1955), Some statistical problems in relating experimental data to predicting performance of a production process, *J. Amer. Statist. Ass.* **50,** 163.

Андерсон (Anderson T. W.) (1960), A modification of the sequential probability ratio test to reduce sample size, *Ann. Math. Statist.* **31,** 165.

Андерсон (Anderson T. W.) (1962), Least squares and best unbiased estimates, *Ann. Math. Statist.* **33,** 266.

Андерсон и Дарлинг (Anderson T. W. and Darling D. A.) (1952), Asymptotic theory of certain «goodness of fit» criteria based on stochastic processes, *Ann. Math. Statist.* **23,** 193.

Андерсон и Дарлинг (Anderson T. W. and Darling D. A.) (1954), A test of goodness of fit, *J. Amer. Statist. Ass.* **49,** 765.

Армитедж (Armitage P.) (1947), Some sequential tests of Student's hypothesis, *J. Roy. Statist. Soc.* **B, 9,** 250.

Армитедж (Armitage P.) (1955), Tests for linear trends in proportions and frequencies, *Biometrics* **11,** 375.

Армитедж (Armitage P.) (1957), Restricted sequential procedures, *Biometrika* **44,** 9.

Армитедж (Armitage P.) (1966), The chi-square test for heterogeneity of proportions, after adjustment for stratification, *J. Roy. Statist. Soc.* **B, 28,** 150 (Addendum: **29,** 197).

Армсен (Armsen P.) (1955), Tables for significance tests of 2×2 contingency tables, *Biometrika* **42,** 494.

Арнольд (Arnold H. J.) (1965), Small sample power of the one sample Wilcoxon tests for non-normal shift alternatives, *Ann. Math. Statist.* **36,** 1767.

Асковиц (Askovitz S. I.) (1957), A short-cut graphic method for fitting the best straight line to a series of points according to the criterion of least squares, *J. Amer. Statist. Ass.* **52,** 13.

Аспин (Aspin A. A.) (1948), An examination and further development of a formula arising in the problem of comparing two mean values, *Biometrika* **35,** 88.

Аспин (Aspin A. A.) (1949), Tables for use in comparisons whose accuracy involves two variances, separately estimated, *Biometrika* **36,** 290.

Балмер (Bulmer M. G.) (1958), Confidence limits for distance in the analysis of variance, *Biometrika* **45,** 360.

Баранкин (Barankin E. W.) (1949), Locally best unbiased estimates, *Ann. Math. Statist.* **20,** 477.

Баранкин (Barankin E. W.) (1960a), Application to exponential families of the solution of the minimal dimensionality problem for sufficient statistics, *Bull. Int. Statist. Inst.* **38 (4)**, 141.

Баранкин (Barankin E. W.) (1960b), Sufficient parameters: solution of the minimal dimensionality problem, *Ann. Inst. Statist. Math.* **12**, 91.

Баранкин (Barankin E. W.) (1961), A note on functional minimality of sufficient statistics, *Sankhyā* **A, 23**, 401.

Баранкин и Катц (Barankin E. W. and Katz M., Jr.) (1959), Sufficient statistics of minimal dimension, *Sankhyā* **21**, 217.

Баранкин и Майтра (Barankin E. W. and Maitra A. P.) (1963), Generalization of the Fisher — Darmois — Koopman — Pitman theorem on sufficient statistics, *Sankhyā* **A, 25**, 217.

Барнард (Barnard G. A.) (1946), Sequential tests in industrial statistics, *Suppl. J. Roy. Statist. Soc.* **8**, 1.

Барнард (Barnard G. A.) (1947a), Significance tests for 2×2 tables, *Biometrika* **34**, 123.

Барнард (Barnard G. A.) (1947b), 2×2 tables. A note on E. S. Pearson's paper, *Biometrika* **34**, 168.

Барнард (Barnard G. A.) (1950), On the Fischer — Behrens test, *Biometrika* **37**, 203.

Барнард (Barnard G. A.) (1963), The logic of least squares, *J. Roy. Statist. Soc.* **B, 25**, 124.

Барнетт (Barnett V. D.) (1966a), Evaluation of the maximum-likelihood estimator where the likelihood equation has multiple roots, *Biometrika* **53**, 151.

Барнетт (Barnett V. D.) (1966b), Order statistics estimators of the location of the Cauchy distribution, *J. Amer. Statist. Ass.* **61**, 1205.

Барр (Barr D. R.) (1966), On testing the equality of uniform and related distributions, *J. Amer. Statist. Ass.* **61**, 856.

Бартки (Bartky W.) (1943), Multiple sampling with constant probability, *Ann. Math. Statist.* **14**, 363.

Бартлетт (Bartlett M. S.) (1935a), The effect of non-normality on the t-distribution, *Proc. Camb. Phil. Soc.* **31**, 223.

Бартлетт (Bartlett M. S.) (1935b), Contingency table interactions, *Suppl. J. Roy. Statist. Soc.* **2**, 248.

Бартлетт (Bartlett M. S.) (1936), The information available in small samples, *Proc. Camb. Phil. Soc.* **32**, 560.

Бартлетт (Bartlett M. S.) (1937), Properties of sufficiency and statistical tests, *Proc. Roy. Soc.* **A, 160**, 268.

Бартлетт (Bartlett M. S.) (1938), The characteristic function of a conditional statistic, *J. Lond. Math. Soc.* **13**, 62.

Бартлетт (Bartlett M. S.) (1939), A note on the interpretation of quasi-sufficiency, *Biometrika* **31**, 391.

Бартлетт (Bartlett M. S.) (1949), Fitting a straight line when both variables are subject to error, *Biometrics* **5**, 207.

Бартлетт (Bartlett M. S.) (1951), An inverse matrix adjustment arising in discriminant analysis, *Ann. Math. Statist.* **22**, 107.

Бартлетт (Bartlett M. S.) (1953, 1955), Approximate confidence intervals, *Biometrika* **40**, 12, 306 and **42**, 201.

Бартоломью (Bartholomew D. J.) (1961), Ordered tests in the analysis of variance, *Biometrika* **48**, 325.

Бартон (Barton D. E.) (1953), On Neyman's smooth test of goodness of fit and its power with respect to a particular system of alternatives, *Skand. Aktuartidskr.* **36**, 24.

Бартон (Barton D. E.) (1955), A form of Neyman's ψ_k^2 test of goodness of fit applicable to grouped and discrete data, *Skand. Aktuartidskr.* **38**, 1.

Бартон (Barton D. E.) (1956), Neyman's ψ_k^2 tests of goodness of fit when the null hypothesis is composite, *Skand. Aktuartidskr.* **39,** 216.

Басу (Basu A. P.) (1965), On some tests of hypotheses relating to the exponential distribution when some outliers are present. *J. Amer. Statist. Ass.* **60,** 548.

Басу Д. (Basu D.) (1955), On statistics independent of a complete sufficient statistic, *Sankhyā* **15,** 377.

Бахадур (Bahadur R. R.) (1958), Examples of inconsistency of maximum likelihood estimates, *Sankhyā* **20,** 207.

Бахадур (Bahadur R. R.) (1964), On Fisher's bound for asymptotic variances, *Ann. Math. Statist.* **35,** 1545.

Бейтмэн (Bateman G. I.) (1949), The characteristic function of a weighted sum of non-central squares of normal variables subject to s linear restraints, *Biometrika* **36,** 460.

Белл и Доксум (Bell C. B. and Doksum K. A.) (1965), Some new distribution-free statistics, *Ann. Math. Statist.* **36,** 203.

Беннетт и Сюй (Bennett B. M. and Hsu P.) (1960), On the power function of the exact test for the 2×2 contingency table, *Biometrika* **47,** 393.

Бенсон (Benson F.) (1949), A note on the estimation of mean and standard deviation from quantiles, *J. Roy. Statist. Soc.* **B, 11,** 91.

Берджер (Berger A.) (1961), On comparing intensities of association between two binary characteristics in two different populations, *J. Amer. Statist. Ass.* **56,** 889.

Берксон (Berkson J.) (1938), Some difficulties of interpretation encountered in the application of the chi-square test, *J. Amer. Statist. Ass.* **33,** 526.

Берксон (Berkson J.) (1950), Are there two regressions? *J. Amer. Statist. Ass.* **45,** 164.

Берксон (Berkson J.) (1955), Maximum likelihood and minimum χ^2 estimates of the logistic function, *J. Amer. Statist. Ass.* **50,** 130.

Берксон (Berkson J.) (1956), Estimation by least squares and by maximum likelihood, *Proc. 3rd Berkeley Symp. Math. Statist. and Prob.* **1,** 1.

Бёрмэн (Burman J. R.) (1946), Sequential sampling formulae for a binomial population, *J. Roy. Statist. Soc.* **B,** 8, 98.

Бернштейн С. Н. (1928), Fondements géométriques de la théorie des corrélations, *Metron* **7** (2), 3.

Бёрч (Birch M. W.) (1963), Maximum likelihood in three-way contingency tables, *J. Roy. Statist. Soc.* **B, 25,** 220.

Бёрч (Birch M. W.) (1964a), A note on the maximum likelihood estimation of a linear structural relationship, *J. Amer. Statist. Ass.* **59,** 1175.

Бёрч (Birch M. W.) (1964b), A new proof of the Pearson — Fisher theorem, *Ann. Math. Statist.* **35,** 817.

Бёрч (Birch M. W.) (1964c), The detection of partial association, I: The 2×2 case, *J. Roy. Statist. Soc.* **B, 26,** 313.

Бёрч (Birch M. W.) (1965), The detection of partial association, II: The general case, *J. Roy. Statist. Soc.* **B, 27,** 111.

Бикел (Bickel P. J.) (1965), On some robust estimates of location, *Ann. Math. Statist.* **36,** 847.

Бирнбаум (Birnbaum Z. W.) (1952), Numerical tabulation of the distribution of Kolmogorov's statistic for finite sample size, *J. Amer Statist. Ass.* **47,** 425.

Бирнбаум (Birnbaum Z. W.) (1953), On the power of a one-sided test of fit for continuous probability functions, *Ann. Math. Statist.* **24,** 484.

Бирнбаум и Тинджи (Birnbaum Z. W. and Tingey F. H.) (1951), One-sided confidence contours for probability distribution functions, *Ann. Math. Statist.* **22,** 592.

Бирнбаум А. (Birnbaum A.) (1962), On the foundations of statistical inference, *J. Amer. Statist. Ass.* **57**, 269.
Бирнбаум А. и Дадмэн (Birnbaum A. and Dudman J.) (1963), Logistic order statistics, *Ann. Math. Statist.* **34**, 658.
Бирнбаум А. и Хили (Birnbaum A. and Healy W. C., Jr.) (1960), Estimates with prescribed variance based on two-stage sampling, *Ann. Math. Statist.* **31**, 662.
Блейлок (Blalock H. M., Jr.) (1958), Probabilistic interpretations for the mean square contingency, *J. Amer. Statist. Ass.* **53**, 102.
Блекуэлл (Blackwell D.) (1946), On an equation of Wald, *Ann. Math. Statist.* **17**, 84.
Блекуэлл (Blackwell D.) (1947), Conditional expectation and unbiased sequential estimation, *Ann. Math. Statist.* **18**, 105.
Блекуэлл и Гиршик (Blackwell D. and Girshick M. A.) (1954), *Theory of Games and Statistical Decisions*, Wiley, New York. (Русский перевод: Д. Блекуэлл и М. А. Гиршик, Теория игр и статистических решений, ИЛ, М., 1958.)
Блисс, Кокрэн и Тьюки (Bliss C. I., Cochran W. G. and Tukey J. W.) (1956), A rejection criterion based upon the range, *Biometrika* **43**, 448.
Блит и Хатчинсон (Blyth C. R. and Hutchinson D. W.) (1960), Table of Neyman-shortest unbiased confidence intervals for the binomial parameter, *Biometrika* **47**, 381.
Блит и Хатчинсон (Blyth C. R. and Hutchinson D. W.) (1961), Table of Neyman-shortest unbiased confidence intervals for the Poisson parameter, *Biometrika* **48**, 191.
Блом (Blom G.) (1958), *Statistical estimates and transformed Beta-variables*, Almqvist and Wiksell, Stockholm; Wiley, New York.
Бломквист (Blomqvist N.) (1950), On a measure of dependence between two random variables, *Ann. Math. Statist.* **21**, 593.
Блоч (Bloch D.) (1966), A note on the estimation of the location parameter of the Cauchy distribution, *J. Amer. Statist. Ass.* **61**, 852.
Блюм, Кифер и Розенблатт (Blum J. R., Kiefer J. and Rosenblatt M.) (1961), Distribution free tests of independence based on the sample distribution function, *Ann. Math. Statist.* **32**, 485.
Бокс (Box G. E. P.) (1949), A general distribution theory for a class of likelihood criteria, *Biometrika* **36**, 317.
Бокс (Box G. E. P.) (1953), Non-normality and tests on variances, *Biometrika* **40**, 318.
Бокс и Андерсен (Box G. E. P. and Andersen S. L.) (1955), Permutation theory in the derivation of robust criteria and the study of departures from assumption, *J. Roy. Statist. Soc.* **B, 17**, 1.
Боукер (Bowker A. H.) (1946), Computation of factors for tolerance limits on a normal distribution when the sample is large, *Ann. Math. Statist.* **17**, 238.
Боукер (Bowker A. H.) (1947), Tolerance limits for normal distributions, *Selected Techniques of Statistical Analysis*, McGraw-Hill, New York.
Боукер (Bowker A. H.) (1948), A test for symmetry in contingency tables, *J. Amer. Statist. Ass.* **43**, 572.
Браун и Муд (Brown G. W. and Mood A. M.) (1951), On median tests for linear hypotheses, *Proc. 2nd Berkeley Symp. Math. Statist. and Prob.*, 159.
Браун Л. (Brown L.) (1964), Sufficient statistics in the case of independent random variables, *Ann. Math. Statist.* **35**, 1456.
Браун Р. (Brown R. L.) (1957), Bivariate structural relation, *Biometrika* **44**, 84.
Браун Р. и Фередей (Brown R. L. and Fereday F.) (1958), Multivariate linear structural relations, *Biometrika* **45**, 136.

Бриллинджер (Brillinger D. R.) (1962) Examples bearing on the definition of fiducial probability with a bibliography, *Ann. Math. Statist.* **33**, 1349.

Бриллинджер (Brillinger D. R.) (1964), The asymptotic behaviour of Tukey's general method of setting approximate confidence limits (the jackknife) when applied to maximum likelihood estimates, *Rev. Int. Statist. Inst.* **32**, 202.

Бриллинджер (Brillinger D. R.) (1966), An extremal property of the conditional expectation, *Biometrika* **53**, 594.

Бросс и Кастен (Bross, I. D. J. and Kasten, E. L.) (1957), Rapid analysis of 2×2 tables, *J. Amer. Statist. Ass.* **52**, 18.

Брэндинер (Brandner F. A.) (1933), A test of the significance of the difference of the correlation coefficients in normal bivariate samples, *Biometrika* **25**, 102.

Бхаттачарджи (Bhattacharjee G. P.) (1965), Effect of non-normality on Stein's two sample test, *Ann. Math. Statist.* **36**, 651.

Бхаттачария (Bhattacharyya A.) (1943), On some sets of sufficient conditions leading to the normal bivariate distribution, *Sankhyā* **6**, 399.

Бхаттачария (Bhattacharyya A.) (1946—7—8), On some analogues of the amount of information and their use in statistical estimation, *Sankhyā* **8**, 1, 201, 315.

Бхучонгкул (Bhuchongkul S.) (1964), A class of nonparametric tests for independence in bivariate populations, *Ann. Math. Statist.* **35**, 138.

Вайсберг и Битти (Weissberg A. and Beatty, G. H.) (1960), Tables of tolerance-limit factors for normal distributions, *Technometrics* **2**, 483.

Вальд (Wald A.) (1940)*), The fitting of straight lines if both variables are subject to error, *Ann. Math. Statist.* **11**, 284.

Вальд (Wald A.) (1941)*), Asymptotically most powerful tests of statistical hypotheses, *Ann. Math. Statist.* **12**, 1.

Вальд (Wald A.) (1942*), On the power function of the analysis of variance test, *Ann. Math. Statist.* **13**, 434.

Вальд (Wald A.) (1943a)*), Tests of statistical hypotheses concerning several parameters when the number of observations is large, *Trans. Amer. Math. Soc.* **54**, 426.

Вальд (Wald A.) (1943b)*), An extension of Wilks' method for setting tolerance limits, *Ann. Math. Statist.* **14**, 45.

Вальд (Wald A.) (1947), *Sequential Analysis*, Wiley, New York. (Русский перевод: А. Вальд, Последовательный анализ, Физматгиз, М., 1960.)

Вальд (Wald A.) (1949), Note on the consistency of the maximum likelihood estimate, *Ann. Math. Statist.* **20**, 595.

Вальд (Wald A.) (1950), *Statistical Decision Functions*, Wiley, New York. (Русский перевод: А. Вальд, Статистические решающие функции, в сб.: Позиционные игры, «Наука», М., 1967.)

Вальд (Wald A.) (1955)*), Testing the difference between the means of two normal populations with unknown standard deviations, *Selected Papers in Statistics and Probability by Abraham Wald*, McGraw-Hill, New York. (Reprint by Stanford Univ. Press.)

Вальд и Волфовиц (Wald A. and Wolfowitz J.) (1939)*), Confidence limits for continuous distribution functions, *Ann. Math. Statist.* **10**, 105.

Вальд и Волфовиц (Wald A. and Wolfowitz J.) (1940)*), On a test whether two samples are from the same population, *Ann. Math. Statist.* **11**, 147.

Вальд и Волфовиц (Wald A. and Wolfowitz J.) (1946)*), Tolerance limits for a normal distribution, *Ann. Math. Statist.* **17**, 208.

*) См. примечание на стр. 877.

Ван дер Ваарт (Van der Vaart H. R.) (1950), Some remarks on the power function of Wilcoxon's test for the problem of two samples I, II, *Proc. Kon. Ned. Akad. Wetensch.* **A, 53,** 494 and 507.

Ван дер Ваарт (Van der Vaart H. R.) (1953), An investigation on the power function of Wilcoxon's two sample test if the underlying distributions are not normal, *Proc. Kon. Ned. Akad. Wetensch.* **A, 56,** 438.

Ван дер Варден (Van der Waerden B. L.) (1952), Order tests for the two-sample problem and their power, *Proc. Kon. Ned. Akad. Wetensch.* **A, 55,** 453.

Ван дер Варден (Van der Waerden B. L.) (1953), Order tests for the two-sample problem, *Proc. Kon. Ned. Akad. Wetensch.* **A, 56,** 303 and 311.

Ван Эден (Van Eeden C.) (1963), The relation between Pitman's asymptotic relative efficiency of two tests and the correlation coefficient between their test statistics, *Ann. Math. Statist.* **34,** 1442.

Ван Эден (Van Eeden C.) (1964), Note on the consistency of some distribution-free tests for dispersion, *J. Amer. Statist. Ass.* **59,** 105.

Ватсон (Watson G. S.) (1957a), Sufficient statistics, similar regions and distribution-free tests, *J. Roy. Statist. Soc.* **B, 19,** 262.

Ватсон (Watson G. S.) (1957b), The χ^2 goodness-of-fit tests for normal distributions, *Biometrika* **44,** 336.

Ватсон (Watson G. S.) (1958), On chi-square goodness-of-fit tests for continuous distributions, *J. Roy. Statist. Soc.* **B, 20,** 44.

Ватсон (Watson G. S.) (1959), Some recent results in chi-square goodness-of-fit tests, *Biometrics* **15,** 440.

Ватсон (Watson G. S.) (1961), Goodness-of-fit tests on a circle, *Biometrika* **48,** 109.

Вилкоксон (Wilcoxon F.) (1945), Individual comparisons by ranking methods, *Biometrics Bull.* **1,** 80.

Вилкоксон (Wilcoxon F.) (1947), Probability tables for individual comparisons by ranking methods, *Biometrics* **3,** 119.

Виллегас (Villegas C.) (1961), Maximum likelihood estimation of a linear functional relationship, *Ann. Math. Statist.* **32,** 1048.

Виллегас (Villegas C.) (1964), Confidence region for a linear relation, *Ann. Math. Statist.* **35,** 780.

Волфовиц (Wolfowitz J.) (1947), The efficiency of sequential estimates and Wald's equation for sequential processes, *Ann. Math. Statist.* **18,** 215.

Волфовиц (Wolfowitz J.) (1949), The power of the classical tests associated with the normal distribution, *Ann. Math. Statist.* **20,** 540.

Габриэл (Gabriel K. R.) (1966), Simultaneous test procedures for multiple comparisons on categorical data, *J. Amer. Statist. Ass.* **61,** 1081.

Гальперин (Halperin M.) (1952a), Maximum likelihood estimation in truncated samples, *Ann. Math. Statist.* **23,** 226.

Гальперин (Halperin M.) (1952b), Estimation in the truncated normal distribution, *J. Amer. Statist. Ass.* **47,** 457.

Гальперин (Halperin M.) (1960), Extension of the Wilcoxon — Mann — Whitney test to samples censored at the same fixed point, *J. Amer. Statist. Ass.* **55,** 125.

Гальперин (Halperin M.) (1961a), Confidence intervals from censored samples, *Ann. Math. Statist.* **32,** 828.

Гальперин (Halperin M.) (1961b), Fitting of straight lines and prediction when both variables are subject to error, *J. Amer. Statist. Ass.* **56,** 657.

Гальперин (Halperin M.) (1963), Confidence interval estimation in non-linear regression, *J. Roy. Statist. Soc.* **B, 25,** 330.

Гальперин (Halperin M.) (1964), Interval estimation of non-linear parametric functions. II, *J. Amer. Statist. Ass.* **59,** 168.

Гальперин, Гринхауз, Корнфилд и Залокар (Halperin M., Greenhouse S. W., Cornfield J. and Zalokar J.) (1955), Tables of percen-

tage points for the studentized maximum absolute deviate in normal samples, *J. Amer. Statist. Ass.* **50,** 185.
Гальперин и Мантел (Halperin M. and Mantel N.) (1963), Interval estimation of non-linear parametric functions, *J. Amer. Statist. Ass.* **58,** 611.
Гамбел (Gumbel E. J.) (1943), On the reliability of the classical chi-square test, *Ann. Math. Statist.* **14,** 253.
Гамильтон (Hamilton M.) (1948), Nomogram for the tetrachoric correlation coefficient, *Psychometrika* **13,** 259.
Гарвуд (Garwood F.) (1936), Fiducial limits for the Poisson distribution, *Biometrika* **28,** 437.
Гастверт (Gastwirth J. L.) (1965), Asymptotically most powerful rank tests for the two-sample problem with censored data, *Ann. Math. Statist.* **36,** 1243.
Гастверт (Gastwirth J. L.) (1966), On robust procedures, *J. Amer. Statist. Ass.* **61,** 929.
Гаттмэн (Guttman I.) (1957), On the power of optimum tolerance regions when sampling from normal distributions, *Ann. Math. Statist.* **28,** 773.
Гафариан (Gafarian A. V.) (1964), Confidence bands in straight line regression, *J. Amer. Statist. Ass.* **59,** 182.
Гейен (Gayen A. K.) (1949), The distribution of «Student's» t in random samples of any size drawn from non-normal universes, *Biometrika* **36,** 353.
Гейен (Gayen A. K.) (1950), The distribution of the variance ratio in random samples of any size drawn from non-normal universes. Significance of difference between the means of two non-normal samples, *Biometrika* **37,** 236, 399.
Гейен (Gayen A. K.) (1951), The frequency distribution of the product-moment correlation coefficient in random samples of any size drawn from non-normal universes, *Biometrika* **38,** 219.
Гейссер и Корнфилд (Geisser S. and Cornfield J.) (1963), Posterior distributions for multivariate normal parameters, *J. Roy. Statist. Soc.* **B, 25,** 368.
Гентер и Уиткомб (Guenther W. C. and Whitcomb M. G.) (1966), Critical regions for tests of interval hypotheses about the variance, *J. Amer. Statist. Ass.* **61,** 204.
Гест (Guest P. G.) (1954), Grouping methods in the fitting of polynomials to equally spaced observations, *Biometrika* **41,** 62.
Гест (Guest P. G.) (1956), Grouping methods in the fitting of polynomials to unequally spaced observations, *Biometrika* **43,** 149.
Гёфдинг (Hoeffding W.) (1948a), A class of statistics with asymptotically normal distribution, *Ann. Math. Statist.* **19,** 293.
Гёфдинг (Hoeffding W.) (1948b), A non-parametric test of independence, *Ann. Math. Statist.* **19,** 546.
Гёфдинг (Hoeffding W.) (1950), «Optimum» non-parametric tests, *Proc. 2nd Berkeley Symp. Math. Statist. and Prob.*, 83.
Гёфдинг (Hoeffding W.) (1952), The large-sample power of tests based on permutations of observations, *Ann. Math. Statist.* **23,** 169.
Гёфдинг (Hoeffding W.) (1953), On the distribution of the expected values of the order statistics, *Ann. Math. Statist.* **24,** 93.
Гёфдинг (Hoeffding W.) (1965), Asymptotically optimal tests for multinomial distributions, *Ann. Math. Statist.* **36,** 369.
Гехан (Gehan E. A.) (1965a), A generalized Wilcoxon test for comparing arbitrarily singly-censored samples, *Biometrika* **52,** 203.
Гехан (Gehan E. A.) (1965b), A generalized two-sample Wilcoxon test for doubly censored data, *Biometrika* **52,** 650.
Гиббонс (Gibbons J. D.) (1964), Effect of non-normality on the power function of the sign test, *J. Amer. Statist. Ass.* **59,** 142.

Гибсон и Джоуэтт (Gibson W. M. and Jowett G. H.) (1957), Three-group regression analysis. Part I, Simple Regression Analysis, *Applied Statist.* **6**, 114.

Гилфорд и Лайонс (Guilford J. P. and Lyons T. C.) (1942), On determining the reliability and significance of a tetrachoric coefficient of correlation, *Psychometrika* **7**, 243.

Гири (Geary R. C.) (1936), The distribution of «Student's» ratio for non-normal samples, *Suppl. J. Roy. Statist. Soc.* **3**, 178.

Гири (Geary R. C.) (1942a), The estimation of many parameters, *J. Roy. Statist. Soc.* **105**, 213.

Гири (Geary R. C.) (1942b), Inherent relations between random variables, *Proc. Roy. Irish Acad.* **A, 47**, 63.

Гири (Geary R. C.) (1943), Relations between statistics; the general and the sampling problem when the samples are large, *Proc. Roy. Irish Acad.* **A, 49**, 177.

Гири (Geary R. C.) (1944), Comparison of the concepts of efficiency and closeness for consistent estimates of a parameter, *Biometrika* **33**, 123.

Гири (Geary R. C.) (1947), Testing for normality, *Biometrika* **34**, 209.

Гири (Geary R. C.) (1949), Determination of linear relations between systematic parts of variables with errors of observation the variances of which are unknown, *Econometrica* **17**, 30.

Гири (Geary R. C.) (1953), Non-linear functional relationship between two variables when one variable is controlled, *J. Amer. Statist. Ass.* **48**, 94.

Гиршик (Girshick M. A.) (1946), Contributions to the theory of sequential analysis, *Ann. Math. Statist.* **17**, 123 and 282.

Гласс (Glass D. V.), ed. (1954), *Social Mobility in Britain*, Routledge and Kegan Paul, London.

Говиндараджулу (Govindarajulu Z.) (1964), A supplement to Mendenhall's bibliography on life testing and related topics, *J. Amer. Statist. Ass.* **59**, 1231.

Говиндараджулу (Govindarajuly Z.) (1966), Best linear estimates under symmetric censoring of the parameters of a double exponential population, *J. Amer. Statist. Ass.* **61**, 248.

Годэмб (Godambe V. P.) (1960), An optimum property of regular maximum likelihood estimation, *Ann. Math. Statist.* **31**, 1208.

Голдбергер (Goldberger A. S.) (1961), Stepwise least squares: residual analysis and specification error, *J. Amer. Statist. Ass.* **56**, 998.

Голдбергер и Джокемс (Goldberger A. S. and Jochems D. V.) (1961), Note on stepwise least squares, *J. Amer. Statist. Ass.* **56**, 105.

Голдмэн и Цайглер (Goldman A. S. and Zeigler R. K.) (1966), Comparisons of some two stage sampling methods, *Ann. Math. Statist*, **37**, 891.

Гохеен и Каврук (Goheen H. W. and Kavruck S.) (1948), A worksheet for tetrachoric *r* and standard error of tetrachoric *r* using Hayes' diagrams and tables, *Psychometrika* **13**, 279.

Граббс (Grubbs F. E.) (1950), Sample criteria for testing outlying observations, *Ann. Math. Statist.* **21**, 27.

Гранди (Grundy P. M.) (1952), The fitting of group truncated and grouped censored normal distributions, *Biometrika* **39**, 252.

Грейбилл и Коннелл (Graybill F. A. and Connell T. L.) (1964), Sample size required for estimating the variance within *d* units of the true value, *Ann. Math. Statist.* **35**, 438.

Грейбилл и Марсалья (Graybill F. A. and Marsaglia G.) (1957), Idempotent matrices and quadratic forms in the general linear hypothesis, *Ann. Math. Statist.* **28**, 678.

Гудмэн (Goodman L. A.) (1963a), On methods for comparing contingency tables, *J. Roy. Statist. Soc.* **A, 126**, 94.

Гудмэн (Goodman L. A.) (1963b), On Plackett's test for contingency table interactions, *J. Roy. Statist. Soc.* **B, 25,** 179.

Гудмэн (Goodman L. A.) (1964a), Simultaneous confidence limits for cross-product ratios in contingency tables, *J. Roy. Statist. Soc.* **B, 26,** 86.

Гудмэн (Goodman L. A.) (1964b), Simultaneous confidence intervals for contrasts among multinomial populations, *Ann. Math. Statist.* **35,** 716.

Гудмэн (Goodman L. A.) (1964c), Interactions in multidimensional contingency tables, *Ann. Math. Statist.* **35,** 632.

Гудмэн (Goodman L. A.) (1964d), Simple methods for analysing three-factor interaction in contingency tables, *J. Amer. Statist. Ass.* **59,** 319.

Гудмэн и Крускал (Goodman L. A. and Kruskal W. H.) (1954, 1959, 1963), Measures of association for cross classifications. Parts I, II and III, *J. Amer. Statist. Ass.* **49,** 732; **54,** 123; **58,** 310.

Гудмэн и Мадански (Goodman L. A. and Madansky A.) (1962), Parameter-free and nonparametric tolerance limits: the exponential case, *Technometrics* **4,** 75.

Гупта и Гнанадесикан (Gupta S. S. and Gnanadesikan M.) (1966), Estimation of the parameters of the logistic distribution, *Biometrika* **53,** 565.

Гупта и Шах (Gupta S. S. and Shah B. K.) (1965), Exact moments and percentage points of the order statistics and the distribution of the range from the logistic distribution, *Ann. Math. Statist.* **36,** 907.

Гхош и Сингх (Ghosh J. K. and Singh R.) (1966), Unbiased estimation of location and scale parameters, *Ann. Math. Statist.* **37,** 1671.

Даниэлс (Daniels H. E.) (1944), The relation between measures of correlation in the universe of sample permutations, *Biometrika* **33,** 129.

Даниэлс (Daniels H. E.) (1948), A property of rank correlations, *Biometrika* **35,** 416.

Даниэлс (Daniels H. E.) (1951—1952), The theory of position finding, *J. Roy. Statist. Soc.* **B, 13,** 186 and **14,** 246.

Даниэлс (Daniels H. E.) (1961), The asymptotic efficiency of a maximum likelihood estimator, *Proc. 4th Berkeley Symp. Math. Statist. and Prob.* **1,** 151.

Даниэлс и Кендалл (Daniels H. E. and Kendall M. G.) (1958), Short proof of Miss Harley's theorem on the correlation coefficient, *Biometrika* **45,** 571.

Данциг (Dantzig G. B.) (1940), On the non-existence of tests of «Student's» hypothesis having power functions independent of σ, *Ann. Math Statist.* **11,** 186.

Дар (Dar S. N.) (1962), On the comparison of the sensitivities of experiments, *J. Roy. Statist. Soc.* **B, 24,** 447.

Дарлинг (Darling D. A.) (1952), On a test for homogeneity and extreme values, *Ann. Math. Statist.* **23,** 450.

Дарлинг (Darling D. A.) (1955), The Cramér — Smirnov test in the parametric case, *Ann. Math. Statist.* **26,** 1.

Дарлинг (Darling D. A.) (1957), The Kolmogorov — Smirnov, Cramér — von Mises tests, *Ann. Math. Statist.* **28,** 823.

Дармуа (Darmois G.) (1935), Sur les lois de probabilité à estimation exhaustive, *C. R. Acad. Sci.*, Paris, **200,** 1265.

Даунтон Ф. (Downton F.) (1953), A note on ordered least-squares estimation, *Biometrika* **40,** 457.

Даунтон (Downton H. F.) (1966), Linear estimates of parameters in the extreme value distribution, *Technometrics* **8,** 3.

Де Гроот (De Groot M. H.) (1959), Unbiased sequential estimation for binomial populations, *Ann. Math. Statist.* **30,** 80.

Дейли (Daly J. F.) (1940), On the unbiased character of likelihood-ratio tests for independence in normal systems, *Ann. Math. Statist.* **11,** 1.

Демпстер (Dempster A. P.) (1963), Further examples of inconsistencies in the fiducial argument, *Ann. Math. Statist.* **34**, 884.
Демпстер (Dempster A. P.) (1964), On the difficulties inherent in Fischer's fiducial argument, *J. Amer. Statist. Ass.* **59**, 56.
Ден Бродер (Den Broeder G. G., Jr.) (1955), On parameter estimation for truncated Pearson Type III distributions, *Ann. Math. Statist.* **26**, 659.
Дёрбин (Durbin J.) (1953), A note on regression when there is extraneous information about one of the coefficients, *J. Amer. Statist. Ass.* **48**, 799.
Дёрбин (Durbin J.) (1954), Errors in variables, *Rev. Int. Statist. Inst.* **22**, 23.
Дёрбин (Durbin J.) (1960), Estimation of parameters in time-series regression models, *J. Roy. Statist. Soc.* **B, 22**, 139.
Дёрбин (Durbin J.) (1961), Some methods of constructing exact tests, *Biometrika* **48**, 41.
Дёрбин и Кендалл (Durbin J. and Kendall M. G.) (1951), The geometry of estimation, *Biometrika* **38**, 150.
Джейкобсон (Jacobson J. E.) (1963), The Wilcoxon two-sample statistic: tables and bibliography, *J. Amer. Statist. Ass.* **58**, 1086.
Дженкинс (Jenkins W. L.) (1955), An improved method for tetrachoric *r*, and Note by J. A. Fishman (1956), *Psychometrika* **20**, 253 and **21**, 305.
Джеффрейс (Jeffreys H.) (1948), *Theory of Probability*, 2nd edn., Oxford Univ. Press.
Джогдео (Jogdeo K.) (1966), On randomized rank score procedures of Bell and Doksum, *Ann. Math. Statist.* **37**, 1697.
Джонкхир (Jonckheere A. R.) (1954), A distribution-free *k*-sample test against ordered alternatives, *Biometrika* **41**, 133.
Джонсон (Johnson N. L.) (1950), On the comparison of estimators, *Biometrika* **37**, 281.
Джонсон (Johnson N. L.) (1959a), On an extension of the connexion between Poisson and χ^2 distributions, *Biometrika* **46**, 352.
Джонсон (Johnson N. L.) (1959b), A proof of Wald's theorem on cumulative sums, *Ann. Math. Statist.* **30**, 1245.
Джонсон (Johnson N. L.) (1961), Sequential analysis: a survey, *J. Roy. Statist. Soc.* **A, 124**, 372.
Джонсон и Уэлч (Johnson N. L. and Welch B. L.) (1939), Applications of the non-central *t* distribution, *Biometrika* **31**, 362.
Джексон (Jackson J. E.) (1960), Bibliography on sequential analysis, *J. Amer. Statist. Ass.* **55**, 516.
Диксон (Dixon W. J.) (1950), Analysis of extreme values, *Ann. Math. Statist.* **21**, 488.
Диксон (Dixon W. J.) (1951), Ratios involving extreme values, *Ann. Math. Statist.* **22**, 68.
Диксон (Dixon W. J.) (1953a), Processing data for outliers, *Biometrics* **9**, 74.
Диксон (Dixon W. J.) (1953b), Power functions of the Sign Test and power efficiency for normal alternatives, *Ann. Math. Statist.* **24**, 467.
Диксон (Dixon W. J.) (1954), Power under normality of several nonparametric tests, *Ann. Math. Statist.* **25**, 610.
Диксон (Dixon W. J.) (1957), Estimates of the mean and standard deviation of a normal population, *Ann. Math. Statist.* **28**, 806.
Диксон (Dixon W. J.) (1960), Simplified estimation from censored normal samples, *Ann. Math. Statist.* **31**, 385.
Диксон и Муд (Dixon W. J. and Mood A. M.) (1946), The statistical sign test, *J. Amer. Statist. Ass.* **41**, 557.
Димер и Вотоу (Deemer W. L., Jr. and Votaw D. F., Jr.) (1955), Estimation of parameters of truncated or censored exponential distributions, *Ann. Math. Statist.* **26**, 498.

Додж и Ромиг (Dodge H. F. and Romig H. G.) (1944), *Sampling Inspection Tables*, Wiley, New York.

Дорфф и Гёрлэнд (Dorff M. and Gurland J.) (1961a), Estimation of the parameters of a linear functional relation, *J. Roy. Statist. Soc.* **B, 23,** 160.

Дорфф и Гёрлэнд (Dorff M. and Gurland J.) (1961b), Small sample behaviour of slope estimators in a linear functional relation, *Biometrics* **17,** 283.

Дункан (Duncan A. J.) (1957), Charts of the 10% and 50% points of the operating characteristic curves for fixed effects analysis of variance F-test, $\alpha = 0.10$ and 0.05, *J. Amer. Statist. Ass.* **52,** 345.

Дынкин Е. Б. (1951), Необходимые и достаточные статистики для семейства распределений вероятностей, *Успехи матем. наук* **6,** 68.

Дэвид (David H. A.) (1956), Revised upper percentage points of the extreme studentized deviate from the sample mean, *Biometrika* **43,** 449.

Дэвид, Хартли и Пирсон (David H. A., Hartley H. O. and Pearson E. S.) (1954), The distribution of the ratio, in a single normal sample, of range to standard deviation, *Biometrika* **41,** 482.

Дэвид и Полсон (David H. A. and Paulson A. S.) (1965), The performance of several tests for outliers, *Biometrika* **52,** 429.

Дэвид С. (David S. T.) (1954), Confidence intervals for parameters in Markov autogressive schemes, *J. Roy. Statist. Soc.* **B, 16,** 195.

Дэвид С., Кендалл и Стьюарт (David S. T., Kendall M. G. and Stuart A.) (1951), Some questions of distribution in the theory of rank correlation, *Biometrika* **38,** 131.

Дэвид Ф. (David F. N.) (1937), A note on unbiased limits for the correlation coefficient, *Biometrika* **29,** 157.

Дэвид Ф. (David F. N.) (1938), *Tables of the Correlation Coefficient*, Cambridge Univ. Press.

Дэвид Ф. (David F. N.) (1939), On Neyman's «smooth» test for goodness of fit, *Biometrika* **31,** 191.

Дэвид Ф. (David F. N.) (1947), A χ^2 «smooth» test for goodness of fit, *Biometrika* **34,** 299.

Дэвид Ф. (David F. N.) (1950), An alternative form of χ^2, *Biometrika* **37,** 448.

Дэвид Ф. и Джонсон (David F. N. and Johnson N. L.) (1948), The probability integral transformation when parameters are estimated from the sample, *Biometrika* **35,** 182.

Дэвид Ф. и Джонсон (David F. N. and Johnson N. L.) (1954), Statistical treatment of censored data. Part I. Fundamental formulae, *Biometrika* **41,** 228.

Дэвид Ф. и Джонсон (David F. N. and Johnson N. L.) (1956), Some tests of significance with ordered variables, *J. Roy. Statist. Soc.* **B, 18,** 1.

Дэвис (Davis R. C.) (1951), On minimum variance in nonregular estimation, *Ann. Math. Statist.* **22,** 43.

Дэрроч (Darroch J. N.) (1962), Interactions in multi-factor contigency tables, *J. Roy. Statist. Soc.* **B, 24,** 251.

Закриссон (Zackrisson U.) (1959), The distribution of «Student's» t in samples from individual non-normal populations, *Publ. Statist. Inst., Univ. Gothenburg*, No. 6.

Зискинд (Zyskind G.) (1963), A note on residual analysis, *J. Amer. Statist. Ass.* **58,** 1125.

Ирвин (Irwin J. O.) (1925), The further theory of Francis Galton's individual difference problem. On a criterion for the rejection of outlying observations, *Biometrika* **17,** 100 and 238.

Ирвин (Irwin J. O.) (1949), A note on the sub-division of χ^2 into components, *Biometrika* **36,** 130.

Ирвин (Irwin J. O.) (1959), On the estimation of the mean of a Poisson distribution from a sample with the zero class missing, *Biometrics* **15**, 329.
Иэйтс (Yates F.) (1934), Contingency tables involving small numbers and the χ^2 test, *Suppl. J. Roy. Statist. Soc.* **1**, 217.
Иэйтс (Yates F.) (1939a), An apparent inconsistency arising from test of significance based on fiducial distributions of unknown parameters, *Proc. Camb. Phil. Soc.* **35**, 579.
Иэйтс (Yates F.) (1939b), Tests of significance of the differences between regression coefficients derived from two sets of correlated variates, *Proc. Roy. Soc. Edin.* **A, 59**, 184.
Иэйтс (Yates F.) (1948), The analysis of contingency tables with groupings based on quantitative characters, *Biometrika* **35**, 176.
Иэйтс (Yates F.) (1955), A note on the application of the combination of probabilities test to a set of 2×2 tables, *Biometrika* **42**, 404.
Кац, Кифер и Волфовиц (Kac M., Kiefer J. and Wolfowitz J.) (1955), On tests of normality and other test of goodness of fit based on distance methods, *Ann. Math. Statist*, **26**, 189.
Кейб (Kabe D. G.) (1963), Extension of Cochran's formulae for addition or omission of variate in multiple regression analysis, *J. Amer. Statist. Ass.* **58**, 527.
Кейл (Kale B. K.) (1961), On the solution of the likelihood equation by iteration processes, *Biometrika* **48**, 452.
Кейл (Kale B. K.) (1962), On the solution of likelihood equations by iteration processes. The multiparametric case, *Biometrika* **49**, 479.
Кейнес (Keynes J. M.) (1911), The principal averages and the laws of error which lead to them, *J. Roy. Statist. Soc.* **74**, 322.
Кейпон (Capon J.) (1961), Asymptotic efficiency of certain locally most powerful rank tests, *Ann. Math. Statist.* **32**, 88.
Кемперман (Kemperman J. H. B.) (1956), Generalized tolerance limits, *Ann. Math. Statist.* **27**, 180.
Кемпторн (Kempthorne O.) (1952), *The Design and Analysis of Experiments*, Wiley, New York.
Кендалл (Kendall M. G.) (1951), Regression, structure and functional relationship. Part I, *Biometrika* **38**, 11.
Кендалл (Kendall M. G.) (1952), Regression, structure and functional relationship. Part II, *Biometrika* **39**, 96.
Кендалл (Kendall M. G.) (1962), *Rank Correlation Methods*, 3rd edn., Griffin, London.
Кимбелл (Kimball A. W.) (1954), Short-cut formulas for the exact partition of χ^2 in contingency tables, *Biometrics* **10**, 452.
Кимбелл Б. (Kimball B. F.) (1946), Sufficient statistical estimation functions for the parameters of the distribution of maximum values, *Ann. Math. Statist.* **17**, 299.
Кифер (Kiefer J.) (1952), On minimum variance estimators, *Ann. Math. Statist.* **23**, 627.
Кифер и Волфовиц (Kiefer J. and Wolfowitz J.) (1956), Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters, *Ann. Math. Statist.* **27**, 887.
Кларк (Clark R. E.) (1953), Percentage points of the incomplete beta function, *J. Amer. Statist. Ass.* **48**, 831.
Клоппер и Пирсон (Clopper C. J. and Pearson E. S.) (1934), The use of confidence or fiducial limits illustrated in the case of the binomial, *Biometrika* **26**, 404.
Клоц (Klotz J.) (1962), Nonparametric test for scale, *Ann. Math. Statist.* **33**, 498.
Клоц (Klotz J.) (1963), Small sample power and efficiency for the one sample Wilcoxon and normal scores tests, *Ann. Math. Statist.* **34**, 624.

Клоц (Klotz J.) (1964), On the normal scores two-samples rank test, *J. Amer. Statist. Ass.* **59**, 652.
Клоц (Klotz J.) (1965), Alternative efficiencies for signed rank tests, *Ann. Math. Statist.* **36**, 1759.
Кокрэн (Cochran W. G.) (1937), The efficiencies of the binomial series tests of significance of a mean and of a correlation coefficient, *J. Roy. Statist. Soc.* **100**, 69.
Кокрэн (Cochran W. G.) (1938), The omission or addition of an independent variate in multiple linear regression, *Suppl. J. Roy. Statist. Soc.* **5**, 171.
Кокрэн (Cochran W. G.) (1950), The comparison of percen tage in matched samples, *Biometrika* **37**, 256.
Кокрэн (Cochran W. G.) (1952), The χ^2 test of goodness of fit, *Ann. Math. Statist.* **23**, 315.
Кокрэн (Cochran W. G.) (1954), Some methods for strengthening the common χ^2 tests, *Biometrics* **10**, 417.
Кокс (Cox C. P.) (1958), A concise derivation of general orthogonal polynomials, *J. Roy. Statist. Soc.* **B, 20**, 406.
Кокс Д. (Cox D. R.) (1952a), Sequential tests for composite hypotheses, *Proc. Camb. Phil. Soc.* **48**, 290.
Кокс Д. (Cox D. R.) (1952b), A note on the sequential estimation of means, *Proc. Camb. Phil. Soc.* **48**, 447.
Кокс Д. (Cox D. R.) (1952c), Estimation by double sampling, *Biometrika* **39**, 217.
Кокс Д. (Cox D. R.) (1956), A note on the theory of quick tests, *Biometrika* **43**, 478.
Кокс Д. (Cox D. R.) (1958a), The regression analysis of binary sequences, *J. Roy. Statist. Soc.* **B, 20**, 215.
Кокс Д. (Cox D. R.) (1958b), Some problems connected with statistical inference, *Ann. Math. Statist.* **29**, 357.
Кокс Д. (Cox D. R.) (1961), Tests of separate families of hypotheses, *Proc. 4th Berkeley Symb. Math. Statist. and Prob.* **1**, 105.
Кокс Д. (Cox D. R.) (1962), Further results on tests of separate families of hypotheses, *J. Roy. Statist. Soc.* **B, 24**, 406.
Кокс Д. (Cox D. R.) (1963), Large sample sequential tests for composite hypotheses, *Sankhyā* **A, 25**, 5.
Кокс Д. (Cox D. R.) (1964), Some applications of exponential order scores, *J. Roy. Statist. Soc.* **B, 26**, 103.
Кокс Д. и Стьюарт (Cox D. R. and Stuart A.) (1955), Some quick sign tests for trend in location and dispersion, *Biometrika* **42**, 80.
Колмогоров А. Н. (1933), Sulla determinazione empirica di una legge di distribuzione, *G. Ist. Ital. Attuari* **4**, 83.
Колоджейчик (Kolodziejczyk S.) (1935), On an important class of statistical hypotheses, *Biometrika* **27**, 161.
Кондо (Kondo T.) (1929), On the standard error of the mean square contingency, *Biometrika* **21**, 376.
Конейн (Konijn H. S.) (1956, 1958), On the power of certain tests for independence in bivariate populations, *Ann. Math. Statist.* **27**, 300 and **29**, 935.
Коэн (Cohen A. C., Jr.) (1950a), Estimating the mean and variance of normal populations from singly truncated and doubly truncated samples, *Ann. Math. Statist.* **21**, 557.
Коэн (Cohen A. C., Jr.) (1950b), Estimating parameters of Pearson Type III populations from truncated samples, *J. Amer. Statist. Ass.* **45**, 411.
Коэн (Cohen A. C., Jr.) (1954), Estimation of the Poisson parameter from truncated samples and from censored samples, *J. Amer. Statist. Ass.* **49**, 158.

Коэн (Cohen A. C., Jr) (1957), On the solution of estimating equations for truncated and censored samples from normal populations, *Biometrika* **44**, 225.

Коэн (Cohen A. C., Jr.) (1960a), Estimating the parameter of a modified Poisson distribution, *J. Amer. Statist. Ass.* **55**, 139.

Коэн (Cohen A. C., Jr.) (1960b), Estimating the parameter in a conditional Poisson distribution, *Biometrics* **16**, 203.

Коэн (Cohen A. C., Jr.) (1960c), Estimation in truncated Poisson distribution when zeros and some ones are missing, *J. Amer. Statist. Ass.* **55**, 342.

Коэн и Вудворд (Cohen A. C., Jr. and Woodward J.) (1953), Tables of Pearson — Lee — Fisher functions of singly truncated normal distributions, *Biometrics* **9**, 489.

Крамер (Cramér H.) (1946), *Mathematical Methods of Statistics*, Princeton Univ. Press. (Русский перевод: Г. Крамер, Математические методы статистики, ИЛ, М., 1947.)

Крамер К. (Cramer K. H.) (1963), Tables for constructing confidence limits on the multiple correlation coefficient, *J. Amer. Statist. Ass.* **58**, 1082.

Крауз (Crouse C. F.) (1964), Note on Mood's test, *Ann. Math. Statist.* **35**, 1825.

Кризи (Creasy M. A.) (1956), Confidence limits for the gradient in the linear functional relationship, *J. Roy. Statist. Soc.* **B, 18**, 65.

Кришнан (Krishnan M.) (1966), Locally unbiased type M test, *J. Roy. Statist. Soc.* **B, 28**, 298.

Кришнан (Krishnan M.) (1967), The moments of a doubly non-central t distributions, *J. Amer. Statist. Ass.*, **62**, 278.

Кроу (Crow E. L.) (1956), Confidence intervals for a proportion, *Biometrika* **43**, 423.

Кроу и Гарднер (Crow E. L. and Gardner R. S.) (1959), Confidence intervals for the expectation of a Poisson variable, *Biometrika* **46**, 441.

Крускал (Kruskal W. H.) (1952), A nonparametric test for the several sample problems, *Ann. Math. Statist.* **23**, 525.

Крускал (Kruskal W. H.) (1957), Historical notes on the Wilcoxon unpaired two-sample test, *J. Amer. Statist. Ass* **52**, 356.

Крускал (Kruskal W. H.) (1958), Ordinal measures of association, *J. Amer. Statist. Ass.* **53**, 814.

Крускал (Kruskal W. H.) (1961), The coordinate-free approach to Gauss — Markov estimation, and its application to missing and extra observations, *Proc. 4th Berkeley Symp. Math. Statist. and. Prob.* **1**, 435.

Крускал и Уоллис (Kruskal W. H. and Wallis, W. A.) (1952, 1953), Use of ranks in one-criterion variance analysis, *J. Amer. Statist. Ass.* **47**, 583 and **48**, 907.

Кудо (Kudo A.) (1956), On the testing of outlying observations, *Sankhyā* **17**, 67.

Куисенберри и Дэвид (Quesenberry C. P. and David H. A.) (1961), Some tests for outliers, *Biometrika* **48**, 379.

Куллдорф (Kulldorff G.) (1963), Estimation of one or two parameters of the exponential distribution on the basis of suitably chosen order statistics, *Ann. Math. Statist.* **34**, 1419.

Купмэн (Koopman B. O.) (1936), On distributions admitting a sufficient statistic, *Trans. Amer. Math. Soc.* **39**, 399.

Кэнуй (Quenouille M. H.) (1948), Partitioning chi-squared, *Unpublished.*

Кэнуй (Quenouille M. H.) (1956), Notes on bias in estimation, *Biometrika* **43**, 353.

Кэнуй (Quenouille M. H.) (1958), *Fundamentals of Statistical Reasoning*, Griffin, London.

Кэнуй (Quenouille M. H.) (1959), Tables of random observations from standard distributions, *Biometrika* **46**, 178.

Кэтти и Гёрлэнд (Katti S. K. and Gurland J.) (1962), Efficiency of certain methods of estimation for the negative binomial and the Neyman type A distributions, *Biometrika* **49**, 215.

Лаха (Laha R. G.) (1956), On a characterization of the stable law with finite expectation, *Ann. Math. Statist.* **27**, 187.

Лача (Latscha R.) (1953), Tests of significance in a 2×2 contingency table: extension of Finney's tables, *Biometrika* **40**, 74.

Левонтин и Праут (Lewontin R. C. and Prout T.) (1956), Estimation of the number of different classes in a population, *Biometrics* **12**, 211.

Ле Кам (Le Cam L.) (1953), On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates, *Univ. Calif. Publ. Statist.* **1**, 277.

Леман (Lehmann E. L.) (1947), On optimum tests of composite hypotheses with one constraint, *Ann. Math. Statist.* **18**, 473.

Леман (Lehmann E. L.) (1949), Some comments on large sample tests, *Proc. 1st Berkeley Symp. Math. Statist. and Prob.*, 451.

Леман (Lehmann E. L.) (1950), Some principles of the theory of testing hypotheses, *Ann. Math. Statist.* **21**, 1.

Леман (Lehmann E. L.) (1951), Consistency and unbiassedness of certain non-parametric tests, *Ann. Math. Statist.* **22**, 165.

Леман (Lehmann E. L.) (1959), *Testing Statistical Hypotheses*, Wiley, New York. (Русский перевод: Э. Леман, Проверка статистических гипотез, «Наука», М., 1964.)

Леман (Lehmann E. L.) (1963), Nonparametric confidence intervals for a shift parameter, *Ann. Math. Statist.* **34**, 1507.

Леман и Стейн (Lehman E. L. and Stein C.) (1948), Most powerful tests of composite hypotheses. I. Normal distributions, *Ann. Math. Statist.* **19**, 495.

Леман и Стейн (Lehmann E. L. and Stein C.) (1949), On the theory of some non-parametric hypotheses, *Ann. Math. Statist.* **20**, 28.

Леман и Стейн (Lehmann E. L. and Stein C.) (1950), Completeness in the sequential case, *Ann. Math. Statist.* **21**, 376.

Леман и Шеффе (Lehmann E. L. and Scheffé H.) (1950, 1955), Completeness, similar regions and unbiased estimation, *Sankhyā* **10**, 305 and **15**, 219.

Лемер (Lehmer E.) (1944), Inverse tables of probabilities of errors of the second kind, *Ann. Math. Statist.* **15**, 388.

Линдли (Lindley D. V.) (1947), Regression lines and the linear functional relationship, *Suppl. J. Roy. Statist. Soc.* **9**, 218.

Линдли (Lindley D. V.) (1950), Grouping corrections and maximum likelihood equations, *Proc. Camb. Phil. Soc.* **46**, 106.

Линдли (Lindley D. V.) (1953a), Statistical inference, *J. Roy. Statist. Soc.* **B, 15**, 30.

Линдли (Lindley D. V.) (1953b), Estimation of a functional relationship, *Biometrika* **40**, 47.

Линдли (Lindley D. V.) (1958a), Fiducial distributions and Bayes' theorem, *J. Roy. Statist. Soc.* **B, 20**, 102.

Линдли (Lindley D. V.) (1958b), Discussion of Cox (1958a).

Линдли (Lindley D. V.) (1964), The Bayesian analysis of contingency tables, *Ann. Math. Statist.* **35**, 1622.

Линдли, Ист и Гамильтон (Lindley D. V., East D. A. and Hamilton P. A.) (1960), Tables for making inferences about the variance of a normal distribution, *Biometrika* **47**, 433.

Линник Ю. В. (1964), On the Behrens — Fisher problem, *Bull. Int. Statist. Inst.* **40 (2)**, 833.

Ллойд (Lloyd E. H.) (1952), Least-squares estimation of location and scale parameters using order statistics, *Biometrika* **39**, 88.

Локс, Александер и Байерс (Locks M. O., Alexander M. J. and Byars B. J.) (1963), New tables of the noncentral *t*-distribution, *Aeronaut. Res. Lab.* (*Ohio*), No. ARL 63-19.

Лоли (Lawley D. N.) (1956), A general method for approximating to the distribution of likelihood ratio criteria, *Biometrika* **43**, 295.

Лорент (Laurent A. G.) (1963), Conditional distribution of order statistics and distribution of the reduced *i*th order statistic of the exponential model, *Ann. Math. Statist.* **34**, 652.

Лотон (Lawton W. H.) (1965), Some inequalities for central and non-central distributions, *Ann. Math. Statist.* **36**, 1521.

Льюис (Lewis B. N.) (1962), On the analysis of interaction in multidimensional contingency tables, *J. Roy. Statist. Soc.* **A, 125**, 88.

Льюис П. (Lewis P. A. W.) (1961), Distribution of the Anderson — Darling statistic, *Ann. Math. Statist.* **32**, 1118.

Льюис Т. и Оделл (Lewis T. O. and Odell P. L.) (1966), A generalization of the Gauss — Markov theorem, *J. Amer. Statist. Ass.* **61**, 1063.

Лэнкастер (Lancaster H. O.) (1949a), The combination of probabilities arising from data in discrete distributions, *Biometrika* **36**, 370.

Лэнкастер (Lancaster H. O.) (1949b), The derivation and partition of χ^2 in certain discrete distributions, *Biometrika* **36**, 117.

Лэнкастер (Lancaster H. O.) (1950), The exact partition of χ^2 and its application to the problem of the pooling of small expectations, *Biometrika* **37**, 267.

Лэнкастер (Lancaster H. O.) (1951), Complex contingency tables treated by the partition of χ^2, *J. Roy. Statist. Soc.* **B, 13**, 242.

Лэнкастер (Lancaster H. O.) (1957), Some properties of the bivariate normal distribution considered in the form of a contingency table, *Biometrika* **44**, 289.

Лэнкастер (Lancaster H. O.) (1958), The structure of bivariate distributions, *Ann. Math. Statist.* **29**, 719.

Лэнкастер (Lancaster H. O.) (1960), On tests of independence in several dimensions, *J. Aust. Math. Soc.* **1**, 241.

Лэнкастер (Lancaster H. O.) (1963), Canonical correlations and partitions of χ^2, *Quart. J. Math.*, Oxford **(2)**, **14**, 220.

Лэнкастер и Хэмдэн (Lancaster H. O. and Hamdan M. A.) (1964), Estimation of the correlation coefficient in contingency tables with possibly nonmetrical characters, *Psychometrika* **29**, 383.

Мадански (Madansky A.) (1959), The fitting of straight lines when both variables are subject to error, *J. Amer. Statist. Ass.* **54**, 173.

Мадански (Madansky A.) (1962), More on length of confidence intervals, *J. Amer. Statist. Ass.* **57**, 586.

Мадански (Madansky A.) (1963), Tests of homogeneity for correlated samples, *J. Amer. Statist. Ass.* **58**, 97.

Майерс, Шнейдермэн и Армитедж (Myers M. H., Schneiderman M. A. and Armitage P.) (1966), Boundaries for closed (wedge) sequential *t* test plans, *Biometrika* **53**, 431.

Мак Кей (McKay A. T.) (1935), The distribution of the difference between the extreme observation and the sample mean in samples of *n* from a normal universe, *Biometrika* **27**, 466.

Маккиннон (Mackinnon W. J.) (1964), Table for both the sign test and distribution-free confidence intervals of the median for sample sizes to 1000, *J. Amer. Statist. Ass.* **59**, 935.

Мак Корнак (McCornack R. L.) (1965), Extended tables of the Wilcoxon matched pair signed rank statistic, *J. Amer. Statist. Ass.* **60**, 864.

Мак Стьюарт (MacStewart W.) (1941), A note on the power of the sign test, *Ann. Math. Statist.* **12**, 236.

Манн (Mann H. B.) (1945), Non-parametric test against trend, *Econometrica* **13,** 245.

Манн (Mann H. B.) (1949), *Analysis and Design of Experiments: Analysis of Variance and Analysis of Variance Designs*, Dover, New York.

Манн и Вальд (Mann H. B. and Wald A.) (1942), On the choice of the number of intervals in the application of the chi-square test, *Ann. Math. Statist.* **13,** 306.

Манн и Уитни (Mann H. B. and Whitney D. R.) (1947), On a test of whether one of two random variables is stochastically larger than the other, *Ann. Math. Statist.* **18,** 50.

Мариц (Maritz J. S.) (1953), Estimation of the correlation coefficient in the case of a bivariate normal population when one of the variables is dichotomized, *Psychometrika* **18,** 97.

Марсалья (Marsaglia G.) (1964), Conditional means and covariances of normal variables with singular covariance matrix, *J. Amer. Statist. Ass.* **59,** 1203.

Менденхолл (Mendenhall W.) (1958), A bibliography on life testing and related topics, *Biometrika* **45,** 521.

Менденхолл и Хейдер (Mendenhall W. and Hader R. J.) (1958), Estimation of parameters of mixed exponentially distributed failure time distributions from censored life test data, *Biometrika* **45,** 504.

Мёрфи (Murphy R. B.) (1948), Non-parametric tolerance limits, *Ann. Math. Statist.* **19,** 581.

Микки и Браун (Mickey M. R. and Brown M. B.) (1966), Bounds on the distribution functions of the Behrens — Fisher statistic, *Ann. Math. Statist.* **37,** 639.

Микульский (Mikulski P. W.) (1963), On the efficiency of optimal non-parametric procedures in the two sample case, *Ann. Math. Statist.* **34,** 22.

Миллер Л. (Miller L. H.) (1956), Table of percentage points of Kolmogorov statistics, *J. Amer. Statist. Ass.* **51,** 111.

Миллер (Miller R. G., Jr.) (1964), A trustworthy jackknife, *Ann. Math. Statist.* **35,** 1549.

Мозес (Moses L. E.) (1963), Rank tests of dispersion, *Ann. Math. Statist.* **34,** 973.

Молдон (Mauldon J. G.) (1955), Pivotal quantities for Wishart's and related distributions, and a paradox in fiducial theory, *J. Roy. Statist. Soc.* **B, 17,** 79.

Морэн (Moran P. A. P.) (1950), The distribution of the multiple correlation coefficient, *Proc. Camb. Phil. Soc.* **46,** 521.

Мошмэн (Moshman J.) (1958), A method for selecting the size of the initial sample in Stein's two sample procedure, *Ann. Math. Statist.* **29,** 1271.

Муд (Mood A. M.) (1954), On the asymptotic efficiency of certain nonparametric two-sample tests, *Ann. Math. Statist.* **25,** 514.

Мур П. (Moore P. G.) (1956), The estimation of the mean of a censored normal distribution by ordered variables, *Biometrika* **43,** 482.

Мэсси (Massey F. J., Jr.) (1950a), A note on the estimation of a distribution function by confidence limits, *Ann. Math. Statist.* **21,** 116.

Мэсси (Massey F. J., Jr.) (1950b, 1952), A note on the power of a non-parametric test, *Ann. Math. Statist.* **21,** 440 and **23,** 637.

Мэсси (Massey F. J., Jr.) (1951), The Kolmogorov — Smirnov test of goodness of fit, *J. Amer. Statist. Ass.* **46,** 68.

Найт (Knight W.) (1965), A method of sequential estimation applicable to the hypergeometric, binomial, Poisson, and exponential distributions, *Ann. Math. Statist.* **36,** 1494.

Нарайн (Narain R. D.) (1950), On the completely unbiassed character of tests of independence in multivariate normal systems, *Ann. Math. Statist.* **21,** 293.

Нейман (Neyman J.) (1935), Sur la vérification des hypothèses statistiques composées, *Bull. Soc. Math. France* **63**, 1.

Нейман (Neyman J.) (1937a), «Smooth test» for goodness of fit, *Skand. Aktuartidskr.* **20**, 149.

Нейман (Neyman J.) (1937b), Outline of a theory of statistical estimation based on the classical theory of probability, *Phil. Trans.* **A, 236**, 333.

Нейман (Neyman J.) (1938a), On statistics the distribution of which is independent of the parameters involved in the original probability law of the observed variables, *Statist. Res. Mem.* **2**, 58.

Нейман (Neyman J.) (1938b), Tests of statistical hypotheses which are unbiassed in the limit, *Ann. Math. Statist.* **9**, 69.

Нейман (Neyman J.) (1949), Contribution to the theory of the χ^2 test, *Proc. 1st Berkeley Symp. Math. Statist. and Prob.*, 239.

Нейман (Neyman J.) (1951), Existence of consistent estimates of the directional parameter in a linear structural relation between two variables, *Ann. Math. Statist.* **22**, 497.

Нейман, Ивашкевич и Колоджейчик (Neuman J., Iwaskiewicz K. and Kolodziejczyk S.) (1935), Statistical problems in agricultural experimentation, *Suppl. J. Roy. Statist. Soc.* **2**, 107.

Нейман и Пирсон (Neyman J. and Pearson E. S.) (1928), On the use and interpretation of certain test criteria for the purposes of statistical inference, *Biometrika* **20A**, 175 and 263.

Нейман и Пирсон (Neyman J. and Pearson E. S.) (1931), On the problem of *k* samples, *Bull. Acad. Polon. Sci.* **3**, 460.

Нейман и Пирсон (Neyman J. and Pearson E. S.) (1933a), On the testing of statistical hypotheses in relation to probabilities *a priori*, *Proc. Camb. Phil. Soc.* **29**, 492.

Нейман и Пирсон (Neyman J. and Pearson E. S.) (1933b), On the problem of the most efficient tests of statistical hypotheses, *Phil. Trans.* **A, 231**, 289.

Нейман и Пирсон (Neyman J. and Pearson E. S.) (1936a), Sufficient statistics and uniformly most powerful tests of statistical hypotheses, *Statist. Res. Mem.* **1**, 113.

Нейман и Пирсон (Neyman J. and Pearson E. S.) (1936b), Unbiassed critical regions of Type A and Type A_1, *Statist. Res. Mem.* **1**, 1.

Нейман и Пирсон (Neyman J. and Pearson E. S.) (1938), Certain theorems on unbiassed critical regions of Type A, and Unbiassed tests of simple statistical hypotheses specifying the values of more than one unknown parameter, *Statist. Res. Mem.* **2**, 25.

Нейман и Скотт (Neyman J. and Scott E. L.) (1948), Consistent estimates based on partially consistent observations, *Econometrica* **16**, 1.

Нейман и Скотт (Neyman J. and Scott E. L.) (1951), On certain methods of estimating the linear structural relation, *Ann. Math. Statist.* **22**, 352.

Нейман и Токарска (Neyman J. and Tokarska B.) (1936), Errors of the second kind in testing «Student's» hypothesis, *J. Amer. Statist. Ass.* **31**, 318.

Нётер (Noether G. E.) (1955), On a theorem of Pitman, *Ann. Math. Statist.* **26**, 64.

Нётер (Noether G. E.) (1957), Two confidence intervals for the ratio of two probabilities, and some measures of effectiveness, *J. Amer. Statist. Ass.* **52**, 36.

Нётер (Noether G. E.) (1963), Note on the Kolmogorov statistic in the discrete case, *Metrika* **7**, 115.

Нётер (Noether G. E.) (1967), Wilcoxon confidence intervals for location parameters in the discrete case, *J. Amer. Statist. Ass.* **62**, 184.

Нортон (Norton H. W.) (1945), Calculation of chi-square for complex contingency tables, *J. Amer. Statist. Ass.* **40**, 125.
Нэйр (Nair K. R.) (1940), Tables of confidence intervals for the median in samples from any continuous population, *Sankhyā* **4**, 551.
Нэйр (Nair K. R.) (1948), The distribution of the extreme deviate from the sample mean and its studentized form, *Biometrika* **35**, 118.
Нэйр (Nair K. R.) (1952), Tables of percentage points of the «Studentized» extreme deviate from the sample mean, *Biometrika* **39**, 189.
Нэйр и Банерджи (Nair K. R. and Banerjee K. S.) (1942), A note on fitting of straight lines if both variables are subject to error, *Sankhyā* **6**, 331.
Нэйр и Шривастава (Nair K. R. and Shrivastava M. P.) (1942), On a simple method of curve fitting, *Sankhyā* **6**, 121.
Огберн (Ogburn W. F.) (1935), Factors in the variation of crime among cities, *J. Amer. Statist. Ass.* **30**, 12.
Олкин и Прэтт (Olkin I. and Pratt J. W.) (1958), Unbiased estimation of certain correlation coefficients, *Ann. Math. Statist.* **29**, 201.
Оуэн (Owen D. B.) (1963), Factors for one-sided tolerance limits and for variables sampling plans, *Sandia Corp. Monogr.* (Off. Techn. Serv., Dept. Commerce, Washington), SCR-607.
Оуэн (Owen D. B.) (1965), The power of Student's *t*-test, *J. Amer. Statist. Ass.* **60**, 320.
Пайк (Pyke R.) (1965), Spacings, *J. Roy. Statist. Soc.* **B**, **27**, 395.
Патил и Шоррок (Patil G. P. and Shorrock R.) (1965), On certain properties of the exponential-type families, *J. Roy. Statist. Soc.* **B**, **27**, 94.
Патнайк (Patnaik R. B.) (1948), The power function of the test for the difference between two proportions in a 2×2 table, *Biometrika* **35**, 157.
Патнайк (Patnaik P. B.) (1949), The non-central χ^2 and *F*-distributions and their applications, *Biometrika* **36**, 202.
Пжиборовский и Виленский (Przyborowski J. and Wilenski M.) (1935), Statistical principles of routine work in testing clover seed for dodder, *Biometrika* **27**, 273.
Пиллаи (Pillai K. C. S.) (1959), Upper percentage points of the extreme studentized deviate from the sample mean, *Biometrika* **46**, 473.
Пиллаи и Тиензо (Pillai K. C. S. and Tienzo B. P.) (1959), On the distribution of the extreme studentized deviate from the sample mean, *Biometrika* **46**, 467.
Пирс (Peers H. W.) (1965), On confidence points and Bayesian probability points in the case of several parameters, *J. Roy. Statist. Soc.* **B**, **27**, 9.
Пирсон (Pearson E. S.) (1926), A further note on the distribution of range in samples taken from a normal population, *Biometrika* **18**, 173.
Пирсон (Pearson E. S.) (1938), The probability integral transformation for testing goodness of fit and combining independent tests of significance, *Biometrika* **30**, 134.
Пирсон (Pearson E. S.) (1947), The choice of statistical tests illustrated in the interpretation of data classed in a 2×2 table, *Biometrika* **34**, 139.
Пирсон и Меррингтон (Pearson E. S. and Merrington M.) (1948), 2×2 tables: the power function of the test on a randomized experiment, *Biometrika* **35**, 331.
Пирсон и Стефенс (Pearson E. S. and Stephens M. A.) (1962), The goodness-of-fit tests based on W_N^2 and U_N^2, *Biometrika* **49**, 397.
Пирсон и Хартли (Pearson E. S. and Hartley H. O.) (1951), Charts of the power function for analysis of variance tests derived from the non-central *F*-distribution, *Biometrika* **38**, 112.
Пирсон и Чандра Секар (Pearson E. S. and Chandra Sekar C.) (1936), The efficiency of statistical tools and a criterion for the rejection of outlying observations, *Biometrika* **28**, 308.

Пирсон К. (Pearson K.) (1900) *), On a criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen in random sampling, *Phil. Mag.* (5), **50**, 157.

Пирсон К. (Pearson K.) (1904) *), On the theory of contingency and its relation to association and normal correlation, *Drapers' Co. Memoirs, Biometric Series*, No. 1, London.

Пирсон К. (Pearson K.) (1909), On a new method for determining correlation between a measured character *A* and a character *B*, of which only the percentage of cases wherein *B* exceeds (or falls short of) a given intensity is recorded for each grade of *A*, *Biometrika* **7**, 96.

Пирсон К. (Pearson K.) (1913), On the probable error of a correlation coefficient as found from a fourfold table, *Biometrika* **9**, 22.

Пирсон К. (Pearson K.) (1915), On the probable error of a coefficient of mean-square contingency, *Biometrika* **10**, 570.

Пирсон К. (Pearson K.) (1916), On the general theory of multiple contingency, with special reference to partial contingency, *Biometrika* **11**, 145.

Пирсон К. (Pearson K.) (1917), On the probable error of biserial η, *Biometrika* **11**, 292.

Питмэн (Pitman E. J. G.) (1936), Sufficient statistics and intrinsic accuracy, *Proc. Camb. Phil. Soc.* **32**, 567.

Питмэн (Pitman E. J. G.) (1937a), Significance tests which may be applied to samples from any population, *Suppl. J. Roy. Statist. Soc.* **4**, 119.

Питмэн (Pitman E. J. G.) (1937b), Significance tests which may be applied to samples from any populations. II. The correlation coefficient test, *Suppl. J. Roy. Statist. Soc.* **4**, 225.

Питмэн (Pitman E. J. G.) (1937c), The «closest» estimates of statistical parameters, *Proc. Camb. Phil. Soc.* **33**, 212.

Питмэн (Pitman E. J. G.) (1939a), A note on normal correlation, *Biometrika* **31**, 9.

Питмэн (Pitman E. J. G.) (1939b), Tests of hypotheses concerning location and scale parameters, *Biometrika* **31**, 200.

Питмэн (Pitman E. J. G.) (1948), *Non-Parametric Statistical Inference*, University of North Carolina Institute of Statistics (mimeographed lecture notes).

Питмэн (Pitman E. J. G.) (1957), Statistics and science, *J. Amer. Statist. Ass.* **52**, 322.

Плэкетт (Plackett R. L.) (1949), A historical note on the method of least squares, *Biometrika* **36**, 458.

Плэкетт (Plackett R. L.) (1950), Some theorems in least squares, *Biometrika* **37**, 149.

Плэкетт (Plackett R. L.) (1953), The truncated Poisson distribution, *Biometrics* **9**, 485.

Плэкетт (Plackett R. L.) (1958), Linear estimation from censored data, *Ann. Math. Statist.* **29**, 131.

Плэкетт (Plackett R. L.) (1962), A note on interactions in contingency tables, *J. Roy. Statist. Soc.* **B, 24**, 162.

Плэкетт (Plackett R. L.) (1964), The continuity correction in 2×2 tables, *Biometrika* **51**, 327.

Полсон (Paulson E.) (1941), On certain likelihood-ratio tests associated with the exponential distribution, *Ann. Math. Statist.* **12**, 301.

Полсон (Paulson E.) (1952), An optimum solution to the k-samples slippage problem for the normal distribution, *Ann. Math. Statist.* **23**, 610.

*) См. примечание на стр. 877.

Поттхофф и Уиттингхилл (Potthoff R. F. and Whittinghill M.) (1966a), Testing for homogeneity. I. The binomial and multinomial distributions, *Biometrika* **53**, 167.

Поттхофф и Уиттингхилл (Potthoff R. F. and Whittinghill M.) (1966b), Testing for homogeneity. II. The Poisson distribution, *Biometrika* **53**, 183.

Прайс (Price R.) (1964), Some non-central *F*-distributions expressed in closed form, *Biometrika* **51**, 107.

Пресс (Press S. J.) (1966), A confidence interval comparison of two test procedures proposed for the Behrens — Fisher problem, *J. Amer. Statist. Ass.* **61**, 454.

Прэтт (Pratt J. W.) (1961), Length of confidence intervals, *J. Amer. Statist. Ass.* **56**, 549.

Прэтт (Pratt J. W.) (1963), Shorter confidence intervals for the mean of a normal distribution with known variance, *Ann. Math. Statist.* **34**, 574.

Прэтт (Pratt J. W.) (1964), Robustness of some procedures for the two-sample location problem, *J. Amer. Statist. Ass.* **59**, 665.

Пури (Puri M. L.) (1964), Asymptotic efficiency of a class of *c*-sample tests, *Ann. Math. Statist.* **35**, 102.

Путтер (Putter J.) (1955), The treatment of ties in non-parametric tests, *Ann. Math. Statist.* **26**, 368.

Пэчерс (Pachares J.) (1959), Table of the upper 10% points of the studentized range, *Biometrika* **46**, 461.

Пэчерс (Pachares J.) (1960), Tables of confidence limits for the binomial distribution, *J. Amer. Statist. Ass.* **55**, 521.

Пэчерс (Pachares J.) (1961), Tables for unbiased tests on the variance of a normal population, *Ann. Math. Statist.* **32**, 84.

Рагхавачари (Raghavachari M.) (1965a), The two-sample scale problem when locations are unknown, *Ann. Math. Statist.* **36**, 1236.

Рагхавачари (Raghavachari M.) (1965b), On the efficiency of the normal scores test relative to the *F*-test, *Ann. Math. Statist.* **36**, 1306.

Радж (Raj D.) (1953), Estimation of the parameters of Type III populations from truncated samples, *J. Amer. Statist. Ass.* **48**, 366.

Райдер (Rider P. R.) (1955), Truncated binomial and negative binomial distributions, *J. Amer. Statist. Ass.* **50**, 877.

Рамачандрамурти (Ramachandramurty P. V.) (1966), On some non-parametric estimates for shift in the Behrens — Fisher situation, *Ann. Math. Statist.* **37**, 593.

Рамачандран (Ramachandran K. V.) (1958), A test of variances, *J. Amer. Statist. Ass.* **53**, 741.

Рао Б. (Rao B. R.) (1958a), On the relative efficiencies of BAN estimates based on doubly truncated and censored samples, *Proc. Nat. Inst. Sci. India* **A, 24**, 366.

Рао Б. (Rao B. R.) (1958b), On an analogue of Cramér — Rao's inequality, *Skand. Aktuartidskr.* **41**, 57.

Рао (Rao C. R.) (1945), Information and accuracy attainable in the estimation of statistical parameters, *Bull. Calcutta Math. Soc.* **37**, 81.

Рао (Rao C. R.) (1947), Minimum variance and the estimation of several parameters, *Proc. Camb. Phil. Soc.* **43**, 280.

Рао (Rao C. R.) (1952), *Advanced Statistical Methods in Biometric Research*, Wiley, New York.

Рао (Rao C. R.) (1957), Theory of the method of estimation by minimum chi-square, *Bull. Int. Statist. Inst.* **35** (2), 25.

Рао (Rao C. R.) (1961), Asymptotic efficiency and limiting information, *Proc. 4th Berkeley Symp. Math. Statist. and Prob.* **1**, 531.

Рао (Rao C. R.) (1962a), Efficient estimates and optimum inference procedures in large samples, *J. Roy. Statist. Soc.* **B, 24**, 46.

Рао (Rao C. R.) (1962b), Apparent anomalies and irregularities in maximum likelihood estimation, *Sankhyā* **A, 24,** 73.
Раштон (Rushton S.) (1951), On least squares fitting of orthonormal polynomials using the Choleski method, *J. Roy. Statist. Soc.* **B, 13,** 92.
Резников (Resnikoff G. J.) (1962), Tables to facilitate the computation of percentage points of the noncentral *t*-distribution, *Ann. Math. Statist.* **33,** 580.
Резников и Либерман (Resnikoff G. J. and Lieberman G. J.) (1957), *Tables of the non-central t-distribution*, Stanford Univ. Press.
Рикер (Ricker W. E.) (1937), The concept of confidence or fiducial limits applied to the Poisson frequency distribution, *J. Amer. Statist. Ass.* **32,** 349.
Роббинс (Robbins H.) (1944), On distribution-free tolerance limits in random sampling, *Ann. Math. Statist.*, **15,** 214.
Роббинс (Robbins H.) (1948), The distribution of Student's *t* when the population means are unequal, *Ann. Math. Statist.* **19,** 406.
Робинсон (Robinson D. E.) (1964), Estimates for the points of intersection of two polynomial regressions, *J. Amer. Statist. Ass.* **59,** 214.
Робсон (Robson D. S.) (1959), A simple method for constructing orthogonal polynomials when the independent variables is unequally spaced, *Biometrics* **15,** 187.
Робсон и Уитлок (Robson D. S. and Whitlock J. H.) (1964), Estimation of a truncation point, *Biometrika* **51,** 33.
Розенталь (Rosenthal I.) (1966), Distribution of the sample version of the measure of association, Gamma, *J. Amer. Statist. Ass.* **61,** 440.
Рой (Roy A. R.) (1956), On χ^2-statistics with variable intervals, Technical Report, Standford University, Statistics Department.
Рой Дж. и Митра (Roy J. and Mitra S. K.) (1957), Unbiassed minimum variance estimation in a class of discrete distributions, *Sankhyā* **18,** 371.
Рой К. (Roy K. P.) (1957), A note on the asymptotic distribution of likelihood ratio, *Bull. Calcutta Statist. Ass.* **7,** 73.
Рой С. (Roy S. N.) (1954), Some further results in simultaneous confidence interval estimation, *Ann. Math. Statist.* **25,** 752.
Рой С. и Бозе (Roy S. N. and Bose R. C.) (1953), Simultaneous confidence interval estimation, *Ann. Math. Statist.* **24,** 513.
Рой С. и Кастенбаум (Roy S. N. and Kastenbaum M. A.) (1956), On the hypothesis of no «interaction» in a multiway contingency table, *Ann. Math. Statist.* **27,** 749.
Рой С. и Митра (Roy S. N. and Mitra S. K.) (1956), An introduction to some non-parametric generalizations of analysis of variance and multivariate analysis, *Biometrika* **43,** 361.
Ротенберг, Фишер Ф. и Тиланус (Rothenberg T. J., Fisher F. M. and Tilanus C. B.) (1964), A note on estimation from a Cauchy sample, *J. Amer. Statist. Ass.* **59,** 460.
Салех (Saleh A. K. M. E.) (1966), Estimation of the parameters of the exponential distribution based on optimum order statistics in censored samples, *Ann. Math. Statist.* **37,** 1717.
Сандрум (Sundrum R. M.) (1953), The power of Wilcoxon's 2-sample test, *J. Roy. Statist. Soc.* **B, 15,** 246.
Сандрум (Sundrum R. M.) (1954), On the relation between estimating efficiency and the power of tests, *Biometrika* **41,** 542.
Санкаран (Sankaran M.) (1964), On an analogue of Bhattacharya bound, *Biometrika* **51,** 268.
Сархан (Sarhan A. E.) (1954), Estimation of the mean and standard deviation by order statistics, *Ann. Math. Statist.* **25,** 317.
Сархан (Sarhan A. E.) (1955), Estimation of the mean and standard deviation by order statistics. II, III, *Ann. Math. Statist.* **26,** 505 and 576.

Сархан и Гринберг (Sarhan A. E. and Greenberg B. G.) (1956), Estimation of location and scale parameters by order statistics from singly and doubly censored samples. I. The normal distribution up to samples of size 10, *Ann. Math. Statist.* **27**, 427.

Сархан и Гринберг (Sarhan A. E. and Greenberg B. G.) (1957), Tables for best linear estimates by order statistics of the parameters of single exponential distributions from singly and doubly censored samples, *J. Amer. Statist. Ass.* **52**, 58.

Сархан и Гринберг (Sarhan A. E. and Greenberg B. G.) (1958), Estimation of location and scale parameters by order statistics from singly and doubly censored samples. II. Tables for the normal distribution for samples of sizes $11 \leqslant n \leqslant 15$, *Ann. Math. Statist.* **29**, 79.

Сархан и Гринберг (Sarhan A. E. and Greenberg B. G.) (1959), Estimation of location and scale parameters for the rectangular population from censored samples, *J. Roy Statist. Soc.* **B, 21**, 356.

Сархан и Гринберг (Sarhan A. E. and Greenberg B. G.) (editors) (1962), *Contributions to Order Statistics*, Wiley, New York.

Сархан, Гринберг и Огава (Sarhan A. E., and Greenberg B. G. and Ogawa J.) (1963), Simplified estimates for the exponential distribution, *Ann. Math. Statist.* **34**, 102.

Свами (Swamy P. S.) (1962a), On the amount of information supplied by censored samples of grouped observations in the estimation of statistical parameters, *Biometrika* **49**, 245.

Свами (Swamy P. S.) (1962b), On the joint efficiency of the estimates of the parameters of normal populations based on singly and doubly truncated samples, *J. Amer. Statist. Ass.* **57**, 46.

Свами (Swamy P. S.) (1963), On the amount of information supplied by truncated samples of grouped observations in the estimation of the parameters of normal populations, *Biometrika* **50**, 207.

Сведи Эйзенхарт (Swed F. S. and Eisenhart C.) (1943), Tables for testing randomness of grouping in a sequence of alternatives, *Ann. Math. Statist.* **14**, 66.

Сиддики (Siddiqui M. M.) (1963), Optimum estimators of the parameters of negative exponential distributions from one or two order statistics, *Ann. Math. Statist.* **34**, 117.

Сиджел и Тьюки (Siegel S. and Tukey J. W.) (1960), A nonparametric sum of ranks procedure for relative spread in unpaired samples, *J. Amer. Statist. Ass.* **55**, 429.

Сил (Seal H. L.) (1948), A note on the χ^2 smooth test, *Biometrika* **35**, 202.

Силбайндер (Seelbinder B. M.) (1953), On Stein's two-stage sampling scheme, *Ann. Math. Statist.* **24**, 640.

Силверстоун (Silverstone H.) (1957), Estimating the logistic curve, *J. Amer. Statist. Ass.* **52**, 567.

Силлито (Sillito G. P.) (1951), Interrelations between certain linear systematic statistics of samples from any continuous population, *Biometrika* 38, 377.

Сичел (Sichel H. S.) (1951, 1952), New methods in the statistical evaluation of mine sampling data, *Trans. Inst. Mining Metallurgy* **61**, 261.

Слэктер (Slakter M. J.) (1966), Comparative validity of the chi-square and two modified chi-square goodness-of-fit tests for small but equal expected frequencies, *Biometrika* 53, 619.

Смирнов Н. В. (1936), Sur la distribution de ω^2, *C. R. Acad. Sci.*, Paris **202**, 449.

Смирнов Н. В. (1939a), Об уклонениях эмпирической кривой распределения, *Матем. сб.* **6** (48), № 1, 3.

Смирнов Н. В. (1939b), Оценка расхождения между эмпирическими кривыми распределения в двух независимых выборках, *Бюллетень МГУ* **2**, № 2, 3.

Смирнов Н. В. (1948), Table for estimating the goodness of fit of empirical distributions, *Ann. Math. Statist.*, **19**, 279.
Смит (Smith C. A. B.) (1951), A test for heterogeneity of proportions, *Ann. Eugen.* **16**, 16.
Смит (Smith C. A. B.) (1952), A simplified heterogeneity test, *Ann. Eugen.* **17**, 35.
Смит Дж. (Smith J. H.) (1947), Estimation of linear functions of cell proportions, *Ann. Math. Statist.* **18**, 231.
Смит К. (Smith K.) (1916), On the «best» values of the constants in frequency distributions, *Biometrika* **11**, 262.
Смит У. (Smith W. L.) (1957), A note on truncation and sufficient statistics, *Ann. Math. Statist.* **28**, 247.
Со (Saw J. G.) (1959), Estimation of the normal population parameters given a singly censored sample, *Biometrika* **46**, 150.
Со (Saw J. G.) (1961), Estimation of the normal population parameters given a type I censored sample, *Biometrika* **48**, 367.
Сопер (Soper H. E.) (1914), On the probable error of the biserial expression for the correlation coefficient, *Biometrika* **10**, 384.
Спрент (Sprent P.) (1966), A generalized least-squares approach to linear functional relationships, *J. Roy. Statist. Soc.* **B**, **28**, 278.
Спротт (Sprott D. A.) (1960), Necessary restrictions for distributions *a posteriori*, *J. Roy. Statist. Soc.* **B**, **22**, 312.
Спротт (Sprott D. A.) (1961), An example of an ancillary statistic and the combination of two samples by Bayes' theorem, *Ann. Math. Statist.* **32**, 616.
Стейн (Stein C.) (1945), A two-sample test for a linear hypothesis whose power is independent of the variance, *Ann. Math. Statist.* **16**, 243.
Стейн (Stein C.) (1946), A note on cumulative sums, *Ann. Math. Statist.* **17**, 498.
Стерн (Sterne T. E.) (1954), Some remarks on confidence or fiducial limits, *Biometrika* **41**, 275.
Стефенс (Stephens M. A.) (1963), The distribution of the goodness-of-fit statistic, U_N^2. I, *Biometrika* **50**, 303.
Стефенс (Stephens M. A.) (1964), The distribution of the goodness-of-fit statistic, U_N^2. II, *Biometrika* **51**, 393.
Стивенс (Stevens W. L.) (1939), Distribution of groups in a sequence of alternatives, *Ann. Eugen.* **9**, 10.
Стивенс (Stevens W. L.) (1950), Fiducial limits of the parameter of a discontinuous distribution, *Biometrika* **37**, 117.
Стьюарт (Stuart A.) (1953), The estimation and comparison of strengths of association in contingency tables, *Biometrika* **40**, 105.
Стьюарт (Stuart A.) (1954a), Too good to be true?, *Applied Statist.* **3**, 29.
Стьюарт (Stuart A.) (1954b), Asymptotic relative efficiencies of distribution-free test of randomness against normal alternatives, *J. Amer. Statist. Ass.* **49**, 147.
Стьюарт (Stuart A.) (1954c), The correlation between variate-values and ranks in samples from a continuous distribution, *Brit. J. Statist. Psychol.* **7**, 37.
Стьюарт (Stuart A.) (1955a), A paradox in statistical estimation, *Biometrika* **42**, 527.
Стьюарт (Stuart A.) (1955b), A test for homogeneity of the marginal distributions in a two-way classification, *Biometrika* **42**, 412.
Стьюарт (Stuart A.) (1956), The efficiencies of tests of randomness against normal regression, *J. Amer. Statist. Ass.* **51**, 285.
Стьюарт (Stuart A.) (1958), Equally correlated variates and the multinormal integral, *J. Roy. Statist. Soc.* **B**, **20**, 273.

Стьюдент («Student») (1927), Errors of routine analysis, *Biometrika* **19**, 151.
Субрахманиам (Subrahmaniam K.) (1965), A note on estimation in the truncated Poisson, *Biometrika* **52**, 279.
Сукхатм (Sukhatme P. V.) (1936), On the analysis of k samples from exponential populations with special reference to the problem of random intervals, *Statist. Res. Mem.* **1**, 94.
Сэвидж (Savage I. R.) (1962), *Bibliography of non-parametric statistics*, Harvard Univ. Press.
Сэвидж (Savage I. R.) (1957), On the independence of test of randomness and other hypotheses, *J. Amer. Statist. Ass.* **52**, 53.
Сэмпфорд (Sampford M. R.) (1955), The truncated negative binomial distribution, *Biometrika* **42**, 58.
Сюй (Hsu P. L.) (1941), Analysis of variance from the power function standpoint, *Biometrika* **32**, 62.
Тагути (Taguti G.) (1958), Tables of tolerance coefficients for normal populations, *Rep. Statist. Appl. Res. (JUSE)* **5**, 73.
Тайкью (Tiku M. L.) (1965a), Laguerre series forms of non-central χ^2 and F distributions, *Biometrika* **52**, 415.
Тайкью (Tiku M. L.) (1965b), Chi-square approximations for the distributions of goodness-of-fit statistics U_N^2 and W_N^2, *Biometrika* **52**, 630.
Тайкью (Tiku M. L.) (1966), A note on approximating the non-central F distribution, *Biometrika* **53**, 606.
Тамура (Tamura R.) (1963), On a modification of certain rank tests, *Ann. Math. Statist.* **34**, 1101.
Тартер и Кларк (Tarter M. E. and Clark V. A.) (1965), Properties of the median and other order statistics of logistic variates, *Ann. Math. Statist.* **36**, 1779.
Тейл (Theil H.) (1950), A rank-invariant method of linear and polynomial regression analysis, *Indag. Math.* **12**, 85 and 173.
Тейл и Ван Изерен (Theil H. and Van Yzeren J.) (1956), On the efficiency of Wald's method of fitting straight lines, *Rev. Int. Statist. Inst.* **24**, 17.
Тейт (Tate R. F.) (1953), On a double inequality of the normal distribution, *Ann. Math. Statist.* **24**, 132.
Тейт (Tate R. F.) (1954), Correlation between a discrete and a continuous variable. Point-biserial correlation, *Ann. Math. Statist.* **25**, 603.
Тейт (Tate R. F.) (1955), The theory of correlation between two continuous variables when one is dichotomized, *Biometrika* **48**, 205.
Тейт (Tate R. F.) (1959), Unbiased estimation: functions of locaion and scale parameters, *Ann. Math. Statist.* **30**, 341.
Тейт и Гоэн (Tate R. F. and Goen R. L.) (1958), Minimum variance unbiassed estimation for the truncated Poisson distribution, *Ann. Math. Statist.* **29**, 755.
Тейт и Клетт (Tate R. F. and Klett G. W.) (1959), Optimal confidence intervals for the variance of a normal distribution, *J. Amer. Statist. Ass.* **54**, 674.
Тейчер (Teicher H.) (1961), Maximum likelihood characterization of distributions, *Ann. Math. Statist.* **32**, 1214.
Терпстра (Terpstra T. J.) (1952), The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking, *Proc. Kon. Ned. Akad. Wetensch.* **A, 55**, 327.
Терри (Terry M. E.) (1952), Some rank order tests which are most powerful against specific parametric alternatives, *Ann. Math. Statist.* **23**, 346.
Тодхантер (Todhunter I.) (1865), *A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace*, Macmillan, London.

Томпсон (Thompson W. R.) (1935), On a criterion for the rejection of observations and the distribution of the ratio of deviation to sample standard deviation, *Ann. Math. Statist.* **6**, 214.

Томпсон (Thompson W. R.) (1936), On confidence ranges for the median and other expectation distributions for populations of unknown distribution form, *Ann. Math. Statist.* **7**, 122.

Точер (Tocher K. D.) (1950), Extension of the Neyman — Pearson theory of tests to discontinuous variates, *Biometrika* **37**, 130.

Трикетт, Уэлч и Джеймс (Trickett W. H., Welch B. L. and James G. S.) (1956), Further critical values for the two-means problem, *Biometrika* **43**, 203.

Тьюки (Tukey J. W.) (1947), Non-parametric estimation. II. Statistically equivalent blocks and tolerance regions — the continuous case, *Ann. Math. Statist.* **18**, 529.

Тьюки (Tukey J. W.) (1948), Non-parametric estimation. III. Statistically equivalent blocks and multivariate tolerance resions — the discontinuous case, *Ann. Math. Statist.* **19**, 30.

Тьюки (Tukey J. W.) (1949), Sufficiency, truncation and selection, *Ann. Math. Statist.* **20**, 309.

Тьюки (Tukey J. W.) (1957), Some examples with fiducial relevance, *Ann. Math. Statist.* **28**, 687.

Тэнг (Tang P. C.) (1938), The power function of the analysis of variance tests with tables and illustrations of their use, *Statist. Res. Mem.* **2**, 126.

Тэтчер (Thatcher A. R.) (1964), Relationships between Bayesian and confidence limits for predictions, *J. Roy. Statist. Soc.* **B, 26**, 176.

Уайз (Wise M. E.) (1963), Multinomial probabilities and the χ^2 and X^2 distributions, *Biometrika* **50**, 145.

Уайз (Wise M. E.) (1964), A complete multinomial distribution compared with the X^2 approximation and an improvement to it, *Biometrika* **51**, 277.

Уайт (White C.) (1952), The use of ranks in a test of significance for comparing two treatments, *Biometrics* **8**, 33.

Уайтекер (Whitaker L.) (1914), On Poisson's law of small numbers, *Biometrika* **10**, 36.

Уикселл (Wicksell S. D.) (1917), The correlation function of Type A, *Medd. Lunds. Astr. Obs.*, Series 2, No. 17.

Уикселл (Wicksell S. D.) (1934), Analytical theory of regression, *Medd. Lunds. Astr. Obs.*, Series 2, No. 69.

Уилк, Гнанадесикан и Хьюэтт (Wilk M. B., Gnanadesikan R. and Huyett M. J.) (1962), Estimation of parameters of the gamma distribution using order statistics, *Biometrika* **49**, 525.

Уилкинсон (Wilkinson G. N.) (1961), Estimation of proportion from zero-truncated binomial data, *Biometrics* **17**, 153.

Уилкс (Wilks S. S.) (1938a), The large-sample distribution of the likelihood ratio for testing composite hypotheses, *Ann. Math. Statist.* **9**, 60.

Уилкс (Wilks S. S.) (1938b), Shortest average confidence intervals from large samples, *Ann. Math. Statist.* **9**, 166.

Уилкс (Wilks S. S.) (1938c), Fiducial distributions in fiducial inference, *Ann. Math. Statist.* **9**, 272.

Уилкс (Wilks S. S.) (1941), Determination of sample sizes for setting tolerance limits, *Ann. Math. Statist.* **12**, 91.

Уилкс (Wilks S. S.) (1942), Statistical prediction with special reference to the problem of tolerance limits, *Ann. Math. Statist.* **13**, 400.

Уилкс (Wilks S. S.) (1948), Order statistics, *Bull. Amer. Math. Soc.* **54**, 6.

Уилкс и Дейли (Wilks S. S. and Daly J. F.) (1939), An optimum property of confidence regions associated with the likelihood function, *Ann. Math. Statist.* **10**, 225.

Уилльямс (Williams C. A., Jr.) (1950), On the choice of the number and width of classes for the chi-square test of goodness of fit, *J. Amer. Statist. Ass.* **45**, 77.
Уилльямс Э. (Williams E. J.) (1952), Use of scores for the analysis of association in contingency tables, *Biometrika* **39**, 274.
Уилльямс Э. (Williams E. J.) (1959), The comparison of regression variables, *J. Roy. Statist. Soc.* **B, 21**, 396.
Уиттинг (Witting H.) (1960), A generalized Pitman efficiency for non-parametric tests, *Ann. Math. Statist.* **31**, 405.
Уишарт (Wishart J.) (1931), The mean and second moment coefficient of the multiple correlation coefficient in samples from a normal population, *Biometrika* **22**, 353.
Уишарт (Wishart J.) (1932), A note on the distribution of the correlation ratio, *Biometrika* **24**, 441.
Уокер (Walker A. M.) (1963), A note on the asymptotic efficiency of an asymptotically normal estimator sequence, *J. Roy. Statist. Soc.* **B**, **25**, 195.
Уоллес Д. (Wallace D. L.) (1958), Asymptotic approximations to distributions, *Ann. Math. Statist.* **29**, 635.
Уоллес (Wallace T. D.) (1964), Efficiencies for stepwise regressions, *J. Amer. Statist Ass.* **59**, 1179.
Уоллис (Wallis W. A.) (1951), Tolerance intervals for linear regression, *Proc. 2nd Berkeley Symp. Math. Statist. and Prob.*, 43.
Уолш (Walsh J. E.) (1946), On the power function of the sign test for slippage of means, *Ann. Math. Statist.* **17**, 358.
Уолш (Walsh J. E.) (1947), Concerning the effect of intraclass correlation on certain significance tests, *Ann. Math. Statist.* **18**, 88.
Уолш (Walsh J. E.) (1950a), Some estimates and tests based on the r smallest values in a sample, *Ann. Math. Statist.* **21**, 386.
Уолш (Walsh J. E.) (1950b), Some non-parametric tests of whether the largest observations of a set are too large or too small, *Ann. Math. Statist.* **21**, 583.
Уолш (Walsh J. E.) (1962), Distribution-free tolerance intervals for continuous symmetrical populations, *Ann. Math. Statist.* **33**, 1167.
Уоркинг и Хотеллинг (Working H. and Hotelling H.) (1929), The application of the theory of error to the interpretation of trends, *J. Amer. Statist. Ass.* **24** (Suppl.), 73.
Уэлч (Welch B. L.) (1938), The significance of the difference between two means when the population variances are unequal, *Biometrika* **29**, 350.
Уэлч (Welch B. L.) (1939), On confidence limits and sufficiency, with particular reference to parameters of location, *Ann. Math. Statist.* **10**, 58.
Уэлч (Welch B. L.) (1947), The generalization of «Student's» problem when several different population variances are involved, *Biometrika* **34**, 28.
Уэлч (Welch B. L.) (1965), On comparisons between confidence point procedures in the case of a single parameter, *J. Roy. Statist. Soc.* **B**, **27**, 1.
Уэлч и Пирс (Welch B. L. and Peers H. W.) (1963), On formulae for confidence points based on integrals of weighted likelihoods, *J. Roy. Statist. Soc.* **B**, **25**, 318.
Феллер (Feller W.) (1938), Note on regions similar to the sample space, *Statist. Res. Mem.* **2**, 117.
Феллер (Feller W.) (1948), On the Kolmogorov — Smirnov limit theorems for empirical distributions, *Ann. Math. Statist.* **19**, 177.
Фергюсон (Ferguson T. S.) (1961), On the rejection of outliers, *Proc. 4th Berkeley Symp. Math. Statist. and Prob.* **1**, 253.
Фердоорен (Verdooren L. R.) (1963), Extended tables of critical values for Wilcoxon's test statistic, *Biometrika* **50**, 177.
Ферон и Фуржо (Féron R. and Fourgeaud C.) (1952), Quelques propriétés caractéristiques de la loi de Laplace — Gauss, *Publ. Inst. Statist. Paris* **1**, 44.

Фикс (Fix E.) (1949a), Distributions which lead to linear regressions, *Proc. 1st Berkeley Symp. Math. Statist. and Prob.*, 79.
Фикс (Fix E.) (1949b), Tables of noncentral χ^2, *Univ. Calif. Publ. Statist.* **1**, 15.
Фикс и Ходжес (Fix E. and Hodges J. L., Jr.) (1955), Significance probabilities of the Wilcoxon test, *Ann. Math. Statist.* **26**, 301.
Филлер (Fieller E. C.) (1940), The biological standardization of insulin, *Suppl. J. Roy. Statist. Soc.* **7**, 1.
Филлер (Fieller E. C.) (1954), Some problems in interval estimation, *J. Roy. Statist. Soc.* **B, 16**, 175.
Финни (Finney D. J.) (1941), On the distribution of a variate whose logarithm is normally distributed, *Suppl. J. Roy. Statist. Soc.* **7**, 155.
Финни (Finney D. J.) (1948), The Fischer — Yates test of significance in 2×2 contingency tables, *Biometrika* **35**, 145.
Финни (Finney D. J.) (1949a), The truncated binomial distribution, *Ann. Eugen.* **14**, 319.
Финни (Finney D. J.) (1949b), On a method of estimating frequencies, *Biometrika* **36**, 233.
Финни, Лача, Беннетт и Сюй (Finney D. J., Latscha R., Bennett B. M. and Hsu P.) (1963), *Tables for Testing Significance in a* 2×2 *Contingency Table*, Cambridge Univ. Press.
Фишер (Fisher R. A.) (1921a) *), On the mathematical foundations of theoretical statistics, *Phil. Trans.* **A, 222**, 309.
Фишер (Fisher R. A.) (1921b) *), Studies in crop variation. I. An examination of the yield of dressed grain from Broadbalk, *J. Agric. Sci.* **11**, 107.
Фишер (Fisher R. A.) (1921c), On the «probable error» of a coefficient of correlation deduced from a small sample, *Metron* **1** (4), 3.
Фишер (Fisher R. A.) (1922a) *), On the interpretation of chi-square from contingency tables, and the calculation of *P. J. Roy. Statist. Soc.* **85**, 87.
Фишер (Fisher R. A.) (1922b) *), The goodness of fit regression formulae and the distribution of regression coefficients, *J. Roy. Statist. Soc.* **85**, 597.
Фишер (Fisher R. A.) (1924a), The distribution of the partial correlation coefficient, *Metron* **3**, 329.
Фишер (Fisher R. A.) (1924b), The influence of rainfall on the yield of wheat at Rothamsted, *Phil. Trans.* **B, 213**, 89.
Фишер (Fisher R. A.) (1924c) *), The conditions under which χ^2 measures the discrepancy between observation and hypothesis, *J. Roy. Statist. Soc.* **87**, 442.
Фишер (Fisher R. A.) (1925), *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh.
Фишер (Fisher R. A.) (1925) *), Theory of statistical estimation, *Proc. Camb. Phil. Soc.* **22**, 700.
Фишер (Fisher R. A.) (1928a) *), The general sampling distribution of the multiple correlation coefficient, *Proc. Roy. Soc.* **A, 121**, 654.
Фишер (Fisher R. A.) (1928b) *), On a property connecting the χ^2 measure of discrepancy with the method of maximum likelihood, *Atti Congr. Int. Mat., Bologna* **6**, 94.
Фишер (Fisher R. A.) (1935a), *The Design of Experiments*, Oliver and Boyd, Edinburgh.
Фишер (Fisher R. A.) (1935b) *), The fiducial argument in statistical inference, *Ann. Eugen.* **6**, 391.
Фишер (Fisher R. A.) (1939) *), The comparison of samples with possibly unequal variances, *Ann. Eugen.* **9**, 174.

*) См. примечание на стр. 877.

Фишер (Fisher R. A.) (1941) *), The negative binomial distribution, *Ann. Eugen.* **11**, 182.
Фишер (Fisher R. A.) (1956), *Statistical Methods and Scientific Inference*, Oliver and Boyd, Edinburgh.
Фокс (Fox L.) (1950), Practical methods for the solution of linear equations and the inversion of matrices, *J. Roy. Statist. Soc.* **B, 12**, 120.
Фокс и Хейес (Fox L. and Hayes J. G.) (1951), More practical methods for the inversion of matrices, *J. Roy. Statist. Soc.* **B, 13**, 83.
Фокс М. (Fox M.) (1956), Charts of the power of the F-test, *Ann. Math. Statist.* **27**, 484.
Френд, Вейл и Клюнье-Росс (Freund R. J., Vail R. W. and Clunies-Ross C. W.) (1961), Residual analysis, *J. Amer. Statist. Ass.* **56**, 98 (corrigenda: 1005).
Примэн и Холтон (Freeman G. H. and Halton J. H.) (1951), Note on an exact treatment of contingency, goodness of fit and other problems of significance, *Biometrika* **38**, 141.
Фрэзер (Fraser D. A. S.) (1950), Note on the χ^2 smooth test, *Biometrika* **37**, 447.
Фрэзер (Fraser D. A. S.) (1951), Sequentially determined statistically equivalent blocks, *Ann. Math. Statist.* **22**, 372.
Фрэзер (Fraser D. A. S.) (1953), Nonparametric tolerance regions, *Ann. Math. Statist.* **24**, 44.
Фрэзер (Fraser D. A. S.) (1957), *Nonparametric Methods in Statistics*, Wiley, New York.
Фрэзер (Fraser D. A. S.) (1961a), On fiducial inference, *Ann. Math. Statist.* **32**, 661.
Фрэзер (Fraser D. A. S.) (1961b), The fiducial method and invariance, *Biometrika* **48**, 261.
Фрэзер (Fraser D. A. S.) (1962), On the consistency of the fiducial method, *J. Roy. Statist. Soc.* **B, 24**, 425.
Фрэзер (Fraser D. A. S.) (1963), On sufficiency and the exponential family, *J. Roy. Statist. Soc.* **B, 25**, 115.
Фрэзер (Fraser D. A. S.) (1964), Fiducial inference for location and scale parameters, *Biometrika* **51**, 17.
Фрэзер и Гаттмэн (Fraser D. A. S. and Guttman I.) (1956), Tolerance regions, *Ann. Math. Statist.* **27**, 162.
Фрэзер и Уормлитон (Fraser D. A. S. and Wormleighton R.) (1951), Non-parametric estimation. IV, *Ann. Math. Statist.* **22**, 294.
Хадсон (Hudson D. J.) (1966), Fitting segmented curves whose join points have to be estimated, *J. Amer. Statist. Ass.* **61**, 1097.
Халмош и Сэвидж (Halmos P. R. and Savage L. J.) (1949), Application of the Radon — Nikodym theorem to the theory of sufficient statistics, *Ann. Math Statist.* **20**, 225.
Хальд (Hald A.) (1949), Maximum likelihood estimation of the parameters of a normal distribution which is truncated at a known point, *Skand. Aktuartidskr.* **32**, 119.
Хаммерсли (Hammersley J. M.) (1950), On estimating restricted parameters, *J. Roy. Statist. Soc.* **B, 12**, 192.
Харкнесс (Harkness W. L.) (1965), Properties of the extended hypergeometric distribution, *Ann. Math. Statist.* **36**, 938.
Харкнесс и Катц (Harkness W. L. and Katz L.) (1964), Comparison of the power functions for the test of independence in 2×2 contingency tables, *Ann. Math. Statist.* **35**, 1115.
Харли (Harley B. I.) (1956), Some properties of an angular transformation for the correlation coefficient, *Biometrika* **43**, 219.

*) См. примечание на стр. 877.

Харли (Harley B. I.) (1957), Further properties of an angular transformation of the correlation coefficient, *Biometrika* **44**, 273.
Хартер (Harter H. L.) (1960), Tables of range and studentized range, *Ann. Math. Statist.* **31**, 1122.
Хартер (Harter H. L.) (1961a), Expected values of normal order statistics, *Biometrika* **48**, 151.
Хартер (Harter H. L.) (1961b), Estimating the parameters of negative exponential populations from one or two order statistics, *Ann. Math. Statist.* **32**, 1078.
Хартер (Harter H. L.) (1963), Percentage points of the ratio of two ranges and power of the associated test, *Biometrika* **50**, 187.
Хартер (Harter H. L.) (1964), Criteria for best substitute interval estimators, with an application to the normal distribution, *J. Amer. Statist. Ass.* **59**, 1133.
Хартер и Мур (Harter H. L. and Moore A. H.) (1966a), Iterative maximum-likelihood estimation of the parameters of normal populations from singly and doubly censored samples, *Biometrika* **53**, 205.
Хартер и Мур (Harter H. L. and Moore A. H.) (1966b), Local-maximum-likelihood estimation of the parameters of three-parameter lognormal populations from complete and censored samples, *J. Amer. Statist. Ass.* **61**, 842.
Хартли (Hartley H. O.) (1958), Maximum likelihood estimation from incomplete data, *Biometrics* **14**, 174.
Хартли (Hartley H. O.) (1964), Exact confidence regions for the parameters in non-linear regression laws, *Biometrika* **51**, 347.
Хейес (Hayes S. P., Jr.) (1943), Tables of the standard error of the tetrachoric correlation coefficient, *Psychometrika* **8**, 193.
Хейес (Hayes S. P., Jr.) (1946), Diagrams for computing tetrachoric correlation coefficients from percentage difference, *Psychometrika* **11**, 163.
Хёйланд (Høyland A.) (1965), Robustness of the Hodges — Lehmann estimates for shift, *Ann. Math. Statist.* **36**, 174.
Хейнэм и Говиндараджулу (Haynam G. E. and Govindarajulu Z.) (1966), Exact power of Mann — Whitney test for exponential and rectangular alternatives, *Ann. Math. Statist.* **37**, 945.
Хемелрейк (Hemelrijk J.) (1952), A theorem on the Sign Test when ties are present, *Proc. Kon. Ned. Akad. Wetensch.* **A, 55**, 322.
Хеннан Э. (Hannan E. J.) (1956), The asymptotic powers of certain tests based on multiple correlations, *J. Roy. Statist. Soc.* **B, 18**, 227.
Хеннан и Тейт (Hannan J. F. and Tate R. F.) (1965), Estimation of the parameters for a multivariate normal distribution when one variable is dichotomized, *Biometrika* **52**, 664.
Хеннан и Харкнесс (Hannan J. F. and Harkness W. L.) (1963), Normal approximation to the distribution of two independent binomial conditional on fixed sum, *Ann. Math. Statist.* **34**, 1593.
Хили (Healy M. J. R.) (1955), A significance test for the difference in efficiency between two predictors, *J. Roy. Statist. Soc.* **B, 17**, 266.
Хилл (Hill B. M.) (1963a), The three-parameter lognormal distribution and Bayesian analysis of a point-source epidemic, *J. Amer. Statist. Ass.* **58**, 72.
Хилл (Hill B. M.) (1963b), Information for estimating the proportions in mixtures of exponential and normal distributions, *J. Amer. Statist. Ass.* **58**, 918.
Хогбен, Пинкхэм и Уилк (Hogben D., Pinkham R. S. and Wilk M. B.) (1961), The moments of the non-central *t*-distribution, *Biometrika* **48**, 465.
Хогг (Hogg R. V.) (1956), On the distribution of the likelihood ratio, *Ann. Math. Statist.* **27**, 529.
Хогг (Hogg R. V.) (1961), On the resolution of statistical hypotheses, *J. Amer. Statist. Ass.* **56**, 978.

Хогг (Hogg R. V.) (1962), Iterated tests of the equality of several distributions, *J. Amer. Statist. Ass.* **57**, 579.

Хогг и Крэйг (Hogg R. V. and Craig A. T.) (1956), Sufficient statistics in elementary distribution theory, *Sankhyā* **17**, 209.

Хогг и Тейнис (Hogg R. V. and Tanis E. A.) (1963), An iterated procedure for testing the equality of several exponential distributions, *J. Amer. Statist. Ass.* **58**, 435.

Ходжес и Леман (Hodges J. L., Jr. and Lehmann E. L.) (1956), The efficiency of some nonparametric competitors of the *t*-test, *Ann. Math. Statist.* **27**, 324.

Ходжес и Леман (Hodges J. L., Jr. and Lehmann E. L.) (1961), Comparison of the normal scores and Wilcoxon tests, *Proc. 4th Berkeley Symp. Math. Statist. and Prob.* **1**, 307.

Ходжес и Леман (Hodges J. L., Jr. and Lehmann E. L.) (1963), Estimates of location based on rank tests, *Ann. Math. Statist.* **34**, 598.

Холдейн (Haldane J. B. S.) (1939), The mean and variance of χ^2 when used as a test of homogeneity, when expectations are small, *Biometrika* **31**, 346.

Холдейн (Haldane J. B. S.) (1945), On a method of estimating frequencies, *Biometrika* **33**, 222.

Холдейн (Haldane J. B. S.) (1955a), Substitutes for χ^2, *Biometrika* **42**, 265.

Холдейн (Haldane J. B. S.) (1955b), The rapid calculation of χ^2 as a test of homogeneity from a $2 \times n$ table, *Biometrika* **42**, 519.

Холдейн и Смит (Haldane J. B. S. and Smith S. M.) (1956), The sampling distribution of a maximum-likelihood estimate, *Biometrika* **43**, 96.

Холл, Вейсман и Гхош (Hall W. J., Wijsman R. A. and Ghosh J. K.) (1965), The relationship between sufficiency and invariance with applications in sequential analysis, *Ann. Math. Statist.* **36**, 575.

Хора и Бюлер (Hora R. B. and Buehler R. J.) (1966), Fiducial theory and invariant estimation, *Ann. Math. Statist.* **37**, 643.

Хотеллинг (Hotelling H.) (1940), The selection of variates for use in prediction, with some comments on the general problem of nuisance parameters, *Ann. Math. Statist.* **11**, 271.

Хотеллинг (Hotelling H.) (1953), New light on the correlation coefficient and its transforms, *J. Roy. Statist. Soc.* **B, 15**, 193.

Хотеллинг (Hotelling H.) (1961), The behaviour of some standard statistical tests under nonstandard conditions, *Proc. 4th Berkeley Symp. Math. Statist. and Prob.* **1**, 319.

Хотеллинг и Пэбст (Hotelling H. and Pabst M. R.) (1936), Rank correlation and tests on significance involving no assumptions of normality, *Ann. Math. Statist.* **7**, 529.

Хоул (Hoel P. G.) (1951), Confidence regions for linear regression, *Proc. 2nd Berkeley Symp. Math. Statist. and Prob.*, 75.

Хубер (Huber P. J.) (1964), Robust estimation of a location parameter, *Ann. Math. Statist.* **35**, 73.

Хузурбазар (Huzurbazar V. S.) (1948), The likelihood equation, consistency and the maxima of the likelihood function, *Ann. Eugen.* **14**, 185.

Хузурбазар (Huzubazar V. S.) (1949), On a property of distributions admitting sufficient statistics, *Biometrika* **36**, 71.

Хузурбазар (Huzurbazar V. S.) (1955), Confidence intervals for the parameter of a distribution admitting a sufficient statistic when the range depends on the parameter, *J. Roy. Statisti. Soc.* **B, 17**, 86.

Хукер (Hooker R. H.) (1907), Correlation of the weather and crops, *J. Roy. Statist. Soc.* **70**, 1.

Хэджнэл (Hajnal J.) (1961), A two-sample sequential *t*-test, *Biometrika* **48**, 65.

Хэмдэн (Hamdan M. A.) (1963), The number and width of classes in the chi-square test, *J. Amer. Statist. Ass.* **58**, 678.

Хюрениус (Hyrenius H.) (1950), Distribution of «Student» — Fisher's *t* in samples from compound normal functions, *Biometrika* **37**, 429.

Чанда (Chanda K. C.) (1963), On the efficiency of two-sample Mann — Whitney test for discrete populations, *Ann. Math. Statist.* **34**, 612.

Чандлер (Chandler K. N.) (1950), On a theorem concerning the secondary subscripts of deviations in multivariate correlation using Yule's notation, *Biometrika* **37**, 451.

Чепмэн (Chapman D. G.) (1950), Some two sample tests, *Ann. Math. Statist.* **21**, 601.

Чепмэн (Chapman D. G.) (1956), Estimating the parameters of a truncated Gamma distribution, *Ann. Math. Statist.* **27**, 498.

Чепмэн и Роббинс (Chapman D. G. and Robbins H.) (1951), Minimum variance estimation without regularity assumptions, *Ann. Math. Statist.* **22**, 581.

Чернов (Chernoff H.) (1949), Asymptotic studentization in testing of hypotheses, *Ann. Math. Statist.* **20**, 268.

Чернов (Chernoff H.) (1951), A property of some Type A regions, *Ann. Math. Statist.* **22**, 472.

Чернов (Chernoff H.) (1952), A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, *Ann. Math. Statist.* **23**, 493.

Чернов, Гаствирт и Джонс (Chernoff H., Gastwirth J. L. and Johns M. V., Jr.) (1967), Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation, *Ann. Math. Statist.* **38**, 52.

Чернов и Леман (Chernoff H. and Lehmann E. L.) (1954), The use of maximum likelihood estimates in χ^2 tests for goodness of fit, *Ann. Math. Statist.* **25**, 579.

Чернов и Сэвидж (Chernoff H. and Savage I. R.) (1958), Asymptotic normality and efficiency of certain non-parametric test statistics, *Ann. Math. Statist.* **29**, 972.

Чипмэн (Chipman J. S.) (1964), On least squares with insufficient observations, *J. Amer. Statist. Ass.* **59**, 1078.

Чоу, Роббинс и Тейчер (Chow Y. S., Robbins H. and Teicher H.) (1965), Moments of randomly stopped sums, *Ann. Math. Statist*, **36**, 789.

Шапиро и Уилк (Shapiro S. S. and Wilk M. B.) (1965), An analysis of variance test for normality (complete samples), *Biometrika* **52**, 591.

Шах Б. (Shah B. K.) (1966), On the bivariate moments of order statistics from a logistic distribution, *Ann. Math. Statist.* **37**, 1002.

Шах (Shah S. M.) (1961), The asymptotic variances of method of moments estimates of the parameters of the truncated binomial and negative binomial distributions, *J. Amer. Statist. Ass.* **56**, 990.

Шах (Shah S. M.) (1966), On estimating the parameter of a doubly truncated binomial distribution, *J. Amer. Statist. Ass.* **61**, 259.

Шентон (Shenton L. R.) (1949), On the efficiency of the method of moments and Neyman's Type A contagious distribution, *Biometrika* **36**, 450.

Шентон (Shenton L. R.) (1950), Maximum likelihood and the efficiency of the method of moments, *Biometrika* **37**, 111.

Шентон (Shenton L. R.) (1951), Efficiency of the method of moments and the Gram — Charlier Type A distribution, *Biometrika* **38**, 58.

Шентон и Боумэн (Shenton L. R. and Bowman K.) (1963), Higher moments of a maximum-likelihood estimate, *J. Roy. Statist. Soc.* **B, 25**, 305.

Шеффе (Scheffé H.) (1942a), On the theory of testing composite hypotheses with one constraint, *Ann. Math. Statist.* **13**, 280.
Шеффе (Scheffé H.) (1942b), On the ratio of the variances of two normal populations, *Ann. Math. Statist.* **13**, 371.
Шеффе (Scheffé H.) (1943a), On solutions of the Behrens — Fisher problem based on the *t*-distribution, *Ann. Math. Statist.* **14**, 35.
Шеффе (Scheffé H.) (1943b), On a measure problem arising in the theory of non-parametric tests, *Ann. Math. Statist.* **14**, 227.
Шеффе (Scheffé H.) (1944), A note on the Behrens — Fisher problem, *Ann. Math. Statist.* **15**, 430.
Шеффе (Scheffé H.) (1958), Fitting straight lines when one variable is controlled, *J. Amer. Statist. Ass.* **53**, 106.
Шеффе и Тьюки (Scheffé H. and Tukey J. W.) (1945), Non-parametric estimation: I. Validation of order statistics, *Ann. Math. Statist.* **16**, 187.
Шойер и Спёрджен (Scheuer E. M. and Spurgeon R. A.) (1963), Some percentage points of the noncentral *t*-distribution, *J. Amer. Statist. Ass.* **58**, 176.
Эдвардс (Edwards A. W. F.) (1963), The measure of association in a 2×2 table, *J. Roy. Statist. Soc.* **A, 126**, 109.
Эзекиел и Фокс (Ezekiel M. J. B. and Fox K. A.) (1959), *Methods of Correlation and Regression Analysis: Linear and Curvilinear*, Wiley, New York.
Эйзенхарт (Eisenhart C.) (1938), The power function of the χ^2 test, *Bull. Amer. Math. Soc.* **44**, 32.
Эймос (Amos D. E.) (1964), Representation of the central and non-central *t* distributions, *Biometrika* **51**, 451.
Эйткин (Aitken A. C.) (1933), On the graduation of data by the orthogonal polynomials of least squares, *Proc. Roy. Soc. Edin.* **A, 53**, 54.
On fitting polynomials to weighted data by least squares, *Ibid.*, **54**, 1.
On fitting polynomials to data with weighted and correlated errors, *Ibid.*, **54**, 12.
Эйткин (Aitken A. C.) (1935), On least squares and linear combination of observations, *Proc. Roy. Soc. Edin.* **A, 55**, 42.
Эйткин (Aitken A. C.) (1948), On the estimation of many statistical parameters, *Proc Roy. Soc. Edin.* **A, 62**, 369.
Эйткин и Силверстоун (Aitken A. C. and Silverstone H.) (1942), On the estimation of statistical parameters, *Proc. Roy. Soc. Edin.* **A, 61**, 186.
Эллисон (Ellison B. E.) (1964), On two-sided tolerance intervals for a normal distribution, *Ann. Math. Statist.* **35**, 762.
Эль-Бадри и Стефан (El-Badry M. A. and Stephan F. F.) (1955), On adjusting sample tabulations to census counts, *J. Amer. Statist. Ass.* **50**, 738.
Эндрюс (Andrews F. C.) (1954), Asymptotic behaviour of some rank tests for analysis of variance, *Ann. Math. Statist.* **25**, 724.
Энскомб (Anscombe F. J.) (1949a), Tables of sequential inspection schemes to control fraction defective, *J. Roy. Statist. Soc.* **A, 112**, 180.
Энскомб (Anscombe F. J.) (1949b), Large-sample theory of sequential estimation, *Biometrika* **36**, 455.
Энскомб (Anscombe F. J.) (1950), Sampling theory of the negative binomial and logarithmic series distributions, *Biometrika* **37**, 358.
Энскомб (Anscombe F. J.) (1952), Large-sample theory of sequential estimation, *Proc. Camb. Phil. Soc.* **48**, 600.
Энскомб (Anscombe F. J.) (1953), Sequential estimation, *J. Roy. Statist. Soc.* **B, 15**, 1.
Энскомб (Anscombe F. J.) (1960), Rejection of outliers, *Technometrics* **2**, 123.

Энскомб и Пейдж (Anscombe F. J. and Page E. S.) (1954), Sequential tests for binomial and exponential populations, *Biometrika* **41**, 252.
Эпстейн (Epstein B.) (1954), Truncated life tests in the exponential case, *Ann. Math. Statist.* **25**, 555.
Эпстейн и Собел (Epstein B. and Sobel M.) (1953), Life testing, *J. Amer. Statist. Ass.* **48**, 486.
Эпстейн и Собел (Epstein B. and Sobel M.) (1954), Some theorems relevant to life testing from an exponential distribution, *Ann. Math. Statist.* **25**, 373.
Юл (Yule G. U.) (1900), On the association of attributes in statistics, *Phil. Trans.* **A, 194**, 257.
Юл (Yule G. U.) (1907), On the theory of correlation for any number of variables treated by a new system of notation, *Proc. Roy. Soc.* **A, 79**, 182.
Юл (Yule G. U.) (1912), On the methods of measuring association between two attributes, *J. Roy. Statist. Soc.* **75**, 579.
Юл (Yule G. U.) (1926), Why do we sometimes get nonsense-correlations between time-series? — A study in sampling and the nature of time-series, *J. Roy. Statist. Soc.* **89**, 1.
Янг и Пирсон (Young A. W. and Pearson K.) (1915, 1919), On the probable error of a coefficient of contingency without approximation; and *Peccavimus*!, *Biometrika* **11**, 215 and **12**, 259.

ПРИМЕЧАНИЕ. Работы А. Вальда, К. Пирсона и Р. Фишера, отмеченные звездочкой, воспроизведены в следующих сборниках:
R. A. Fischer, *Contributions to Mathematical Statistics*, Wiley, New York, 1950.
Karl Pearson's Early Statistical Papers, Cambridge Univ. Press, London, 1948.
Selected Papers in Statistics and Probability by Abraham Wald, McGraw-Hill, New York, 1955; reprinted by Stanford Univ. Press, 1957.

УКАЗАТЕЛЬ

Александер (Alexander M. J.), таблицы нецентрального t-распределения 342
Аллан (Allan F. E.), ортогональные полиномы 479
Аллен (Allen H. V.), линейность регрессии 553
Альтернатива различия в масштабе (scale-shift alternative) 674
Альтернативная гипотеза (alternative hypothesis) (H_1) см. Гипотезы статистические
Андерсен (Andersen S. L.), устойчивость 622
Андерсон (Anderson T. W.), НК-оценки и МД-оценки 116; контролируемые переменные 545; критерии согласия 604; эффективность ПКОВ 814
АОЭ (ARE) см. Асимптотическая относительная эффективность
Арбетнот (Arbuthnot J.), первое использование критерия знаков 687, сноска
Армитедж (Armitage P.), критерии для тренда в вероятностях 774; дисперсионный критерий X^2 для расслоенной совокупности 776; последовательные критерии 810, 819, 826
Армсен (Armsen P.), таблицы точного критерия независимости в таблицах 2×2 741
Арнольд (Arnold H. J.), мощность критерия симметрии Вилкоксона 681
Арнольд К. (Arnold K. J.), таблицы последовательного t-критерия 819
Асимптотическая относительная эффективность (АОЭ) (asymptotic relative efficiency (ARE)) 355—372; определение 357; АОЭ и производные функции мощности 359—364; критерии, которые не могут быть сравнены с помощью АОЭ 358, 362; АОЭ и максимальная потеря мощности 364—366; АОЭ и эффективность оценивания 367; АОЭ и корреляция 367—368, (упражнение 25.9) 372; случаи не нормальных распределений 368—369.
Асковиц (Askovitz S. I.), графическая подгонка линии регрессии (упражнение 28.19) 497
Аспин (Aspin A. A.), таблицы для задачи о двух средних 202

Байерс (Byars B. J.), таблицы нецентрального t-распределения 342
Байесовские интервалы (Bayesian intervals) 205—207; критическое обсуждение 207—211; б.и. и фидуциальная теория 211—216; б. и. и доверительные интервалы 215
Балмер (Bülmer H. G.), доверительный интервал для параметра «удаленности» 751
Банерджи (Banerjee K. S.), оценивание линейной связи 539
Баранкин (Barankin E. W.), границы для дисперсии 33; достаточность 48; минимальная достаточность 260, сноска
Барнард (Barnard G. A.), оптимальные свойства метода НК 115, сноска; частотное обоснование фидуциального вывода 215; модели для таблиц 2×2 736; последовательные методы 592
Барнетт (Barnett V. D.), итерационное решение уравнения МП 76—77; оценивание параметра сдвига распределения Коши 77
Барр (Barr D. R.) проверка гипотез для нерегулярных распределений 321
Бартки (Bartky W.), обследование с заменой всех дефектных изделий годными 810
Бартлетт (Bartlett M. S.), приближенные доверительные интервалы 156, 175; условное распределение в задаче о двух средних (упражнение 21.10) 218; достаточность и подобные области 254; квазидостаточность 292; аппроксимации распределения статистики ОП 313; модификация статистики ОП (пример 24.4) 316; условная х. ф. 426; поправка в методе НК при появлении дополнительного наблюдения (упражнение 28.20) 497; оценивание функциональной зависимости 538—539, (упражнение 29.11) 555; устойчивость 623—625; взаимодействия в таблицах с несколькими входами 781
Бартоломью (Bartholomew D. J.), эффективность k-выборочных критериев 678
Бартон (Barton D. E.), «гладкие» критерии согласия 596—602
Басу (Basu A. P.), выпадающие наблюдения для экспоненциальных распределений 711
Басу Д. (Basu D.), достаточность, полнота и независимость (упражнение 23.7) 296
Бахадур (Bahadur R. R.), эффективность 70; неединственность МП-оценок (упражнение 18.35) 108
Бейтмэн (Bateman G. I.), нецентральные нормальные величины с ограничениями (упражнение 24.2) 346
Белл (Bell C. B.), критерии, использующие случайные нормальные отклонения 653
Беннетт (Bennett B. M.), критерии независимости в таблицах 2×2 741, 743
Бенсон (Benson F.), оценивание с помощью двух наилучших порядковых статистик (упражнение 32.14) 715
Берджер (Berger A.), сравнение таблиц 2×2 743
Беренс (Behrens W. V.), фидуциальное решение задачи о двух средних 204
Берксон (Berkson J.), выборочные эксперименты для исследования НАН-оценок 136; выбор размера критерия 246; контролируемые переменные 544

Бёрмэн (Burman J. P.), последовательный выбор 807
Бернштейн С. Н., характеризация двумерной нормальности 471 (сноска)
Бёрч (Birch M. W.), МП-оценка структурной зависимости 509, (упражнение 29.18) 557; предельное распределение величины X^2 567; таблицы с несколькими входами 781
Беспорядок (disarray) 640
Бикел (Bickel P. J.) устойчивость и эффективность оценок параметра сдвига 627, 705
Биномиальное распределение (binomial distribution), несмещенное оценивание квадрата параметра θ (пример 17.4) 20; МГД для θ (пример 17.9) 26; оценивание $\theta(1-\theta)$ (пример 17.11) 31; несмещенное оценивание функций от θ (упражнение 17.12) 55; оценивание линейной зависимости между функциями от параметров независимых биномиальных распределений (упражнение 17.25) 57; смещенная МП-оценка при последовательном выборе (упражнение 18.18) 104; эквивалентность МП- и МХК-оценок (упражнение 19.15) 138; доверительные интервалы (пример 20.2) 146, 162—164, (упражнение 20.8) 180; таблицы и графики для доверительных интервалов 162; доверительные интервалы для отношения двух биномиальных параметров (упражнение 20.9) 180; прогноз с помощью байесовского и доверительного методов 215; фидуциальные интервалы (упражнение 21.3) 216; проверка простой гипотезы H_0 о параметре θ (упражнение 22.2) 247, 285; минимальная достаточность (упражнение 23.11) 297; урезание 707; критерий однородности (дисперсионный) 774—775 (упражнения 33.21, 33.22), 787; обратный последовательный выбор (примеры 34.1—34.7) 792—807, (упражнения 34.1—34.3) 827—828; МГД при последовательном оценивании (пример 34.11) 821; двойной выбор (упражнение 34.17), 831; см. Критерий знаков
Бирнбаум (Birnbaum Z. W.), табулирование критерия Колмогорова 611; односторонний критерий согласия 613, 616; вычисление статистики Колмогорова (пример 30.6) 616—617
Бирнбаум А. (Birnbaum A.), принципы условности и правдоподобия 293; логистические порядковые статистики (упражнение 32.12) 714; двойная выборка 826
Бисериальная корреляция (biserial correlation) 411—416, (упражнения 26.5, 26.10—26.12) 418—420
Битти (Beatty G. H.), таблицы толерантных границ для нормального распределения 178
Блейлок (Blalock H. M., Jr.), интерпретация мер X^2 752
Блекуэлл (Blackwell D.), достаточность и МД-оценивание 44; последовательный выбор 806; решающие функции 827
Блисс (Bliss C. I.), выпадающие наблюдения 710
Блит (Blyth C. R.) таблицы доверительных интервалов для биномиального и пуассоновского распределений 165
Блом (Blom G.), методы для логистического распределения (упражнение 18.29) 107; простое оценивание в цензурированных выборках 703; асимптотическая МГД (упражнения 32.8—32.11) 713—714
Бломквист (Blomqvist N.), АОЭ критериев 360; критерий срединной корреляции (упражнение 31.7) 683
Блоч (Bloch D.), оценивание параметра сдвига распределения Коши 77
Блюм (Blum J. R.), распределение критерия Гёфдинга 647
Бозе (Bose R. C.), совместные доверительные интервалы 175
Бокс (Box G. E. P.), аппроксимации распределения статистики ОП 314; устойчивость 622, 626
Боукер (Bowker A. H.), таблицы для толерантных границ 178; аппроксимации толерантных границ (упражнения 20.20—20.22) 182; критерий полной симметрии в таблицах $r \times r$ (упражнение 33.23) 788
Боумэн (Bowman K.), семиинварианты МП-оценок 75
Браун (Brown G. W.), критерий случайности, основанный на медиане (упражнение 31.12) 684
Браун Л. (Brown L.), достаточные статистики 46
Браун М. (Brown M. B.), задача о двух средних 201
Браун Р. (Brown R. L.), данные о функциональной зависимости (пример 29.1) 518; доверительные интервалы для функциональных зависимостей 521—522; линейные функциональные зависимости в случае нескольких переменных 526, (упражнения 29.6—29.8) 554—555
Бриллинджер (Brillinger D. R.), устранение смещения МП-оценок 67; фидуциальные парадоксы 212; максимальная корреляция между метками и порядковыми статистиками 653
Бросс (Bross I. D. J.), графики для критерия независимости в таблицах 2×2 741
Брэндинер (Brandner F. A.), МП-оценки и критерии ОП для параметров корреляции двумерных нормальных распределений (упражнения 26.19—26.22) 421—422
Бхаттачарджи (Bhattacharjee G. P.), отсутствие нормальности и критерий Стейна 826
Бхаттачария (Bhattacharyya A.), нижняя граница для дисперсии 28; ковариация между МГД-оценкой и несмещенными оценками ее семиинвариантов (упражнение 17.4) 54; характеризация двумерной нормальности (упражнения 28.7—28.11) 494
Бхучонгкул (Bhuchongkul S.), критерии независимости 652
Бьенэме — Чебышева неравенство (Bienaymé — Tchebycheff inequality) и состоятельное оценивание (пример 17.2) 17
Бюлер (Buehler R. J.), фидуциальные выводы 212

Вайсберг (Weisberg A.), таблицы толерантных интервалов для нормального распределения 178
Вальд (Wald A.), состоятельность МП-оценок 65; асимптотическая нормальность МП-оценок 83; толерантные интервалы для нормального распределения 176—178; задача о двух средних 202; асимптотическое распределение статистики НК 310; состоятельность критерия 322, 351; асимптотические свойства критерия ОП 330; оптимальное свойство критерия ОП для линейной гипотезы 344; асимптотически наиболее мощные критерии

353; оценивание функциональной зависимости 533—535, (упражнения 29.4—29.5) 554; выбор классов для критерия согласия X^2 576, 583—587; несмещенность критерия согласия X^2 580; односторонний критерий согласия 612; дисперсия статистики X^2 (упражнение 30.3) 618; критерий серий (упражнение 30.8) 619, 673; толерантные области 698; последовательные методы 801—806, 812, 815—816, (упражнения 34.3—34.4, 34.7, 34.11—34.13, 34.18—34.19) 828—831
Ван дер Ваарт (Van der Vaart H. R.), критерий Вилкоксона для двух выборок 664, 667
Ван дер Варден (Van der Waerden B. L.), критерий для двух выборок 674
Ван Изерен (Van Yzeren J.), оценивание функциональной зависимости 539
Ван Эден (Van Eeden C.), состоятельность масштабных критериев 674
Ватсон (Watson G. S.), задача о двух средних (пример 23.10) 268; неполнота и подобные области (упражнение 23.15) 297; распределение критерия согласия X^2 567, (упражнение 30.2) 618; выбор классов для критерия согласия X^2 602; видоизменение критерия Смирнова 605
Вейсман (Wijsman R. A.), инвариантность и достаточность 344, 819
Вероятностное интегральное преобразование (probability integral transformation), в критериях согласия 592, 603, (упражнение 30.10) 620
Взаимодействия (interactions), в таблицах с несколькими входами 781
Взаимозависимость (interdependence) 373—374
Винер (Wiener N.), полнота для параметра сдвига 255
Винзоризованные оценки (winsorized estimators) 704, 710
Виленский (Wilénski H.), доверительные границы для пуассоновского распределения 163
Вилкоксон (Wilcoxon F.), критерий для задачи двух выборок 660
Вилкоксона критерий (Wilcoxon test), 661—667, 673, 674; АОЭ (упражнения 25.3—25.4) 371, 664—667, (упражнения 31.14—31.19, 31.22—31.24) 685—686; для симметрии 681, (упражнение 31.25) 686; обобщение на цензурированные выборки 707; для группы выпадающих наблюдений 712
Виллегас (Villegas C.), линейные функциональные зависимости в случае нескольких переменных 526
Внутриклассовая корреляция (intra-class correlation) 403—407, (упражнение 26.14) 420—421; в многомерном нормальном случае (упражнение 27.17) 459
Волфовиц (Wolfowitz J.), пример, когда МП-оценка не существует (упражнение 18.34) 108; толерантные интервалы для нормального распределения 176—178; состоятельность критерия 322, 351; оптимальные свойства критерия ОП для линейной гипотезы 343; состоятельность МП-оценок 514, 517; односторонний критерий согласия 612; сравнение критерия X^2 и критерия Колмогорова 615; критерии нормальности 618; критерий серий (упражнение 30.8) 619, 674; последовательное оценивание 820—821 (упражнение 34.6) 828
Вотоу (Votaw D. F., Jr.), урезание и цензурирование для экспоненциального распределения 705, (упражнение 32.16) 715
Вспомогательные статистики (ancillary statistics) 292
Вудворд (Woodward J.), урезанное нормальное распределение 703
Выборочная дисперсия (sampling variance), как критерий качества оценки 21; см. Минимальной дисперсии оценки
Выборочная ф. р. (sample d. f.) 602—603
Выпадающие наблюдения (outlying observations) 707—710, (упражнение 32.21) 716—717
Вырожденная матрица (singular matrix), псевдообратная к ней (упражнение 27.21) 460

Габриел (Gabriel K. R.), многократные сравнения в таблицах $r \times c$ 752
Гаек (Hájek J.), задача о двух средних 201
Гальперин (Halperin M.), доверительные интервалы для нелинейных функций 175 (упражнение 28.3) 493; линейная функциональная зависимость 540; МП-оценивание в цензурированных выборках 702, (упражнение 32.15) 715; урезанное нормальное распределение 704; доверительные интервалы для цензурированных выборок 707; критерий Вилкоксона для цензурированных выборок 707; таблицы для стьюдентизованного максимального абсолютного отклонения 710
Гальтон (Galton F.), данные о цвете глаз (пример 33.2) 726
Гамбел (Gumbel E. J.), выбор групп для критерия согласия X^2 576
Гамильтон (Hamilton M.), номограмма для тетрахорической корреляции 410
Гамильтон П. (Hamilton P. A.), таблицы доверительных интервалов для нормальной дисперсии 163
Гамма-распределение (Gamma distribution), МГД-оценивание масштабного параметра (упражнение 17.1) 53; свойства достаточности (упражнение 17.9) 54; оценивание нижней границы (упражнение 17.22) 56; МП-оценивание параметров сдвига и масштаба (пример 18.17) 96; эффективность метода моментов (пример 18.18) 98; доверительные интервалы для масштабного параметра (упражнение 20.1) 179; фидуциальные интервалы для масштабного параметра (пример 21.2) 187; распределение линейной функции от гамма-величин (упражнение 21.8) 217; отсутствие подобных областей (упражнение 23.2) 295; полнота и характеризация (упражнение 23.27) 299; связь с прямоугольным распределением 317—318; АОЭ критерия Вилкоксона (упражнение 31.18) 685; урезание и цензурирование 706
Гарвуд (Garwood F.), таблицы доверительных интервалов для параметра пуассоновского распределения 163
Гарднер (Gardner R. S.), таблицы доверительных границ для параметра пуассоновского распределения 163
Гастверт (Gastwirth J. L.), устойчивость 627; эффективное оценивание с помощью порядковых статистик 703; МД-оценки и ранговые критерии 703; ранговые критерии для цензурированных выборок 707
Гаттмэн (Guttman I.), толерантные интервалы 179

Гаусс (Gauss C.), создатель теории НК 115, 118
Гафариан (Gafarian A. V.), доверительные области для линий регрессии 493
Гейен (Gayen A. K.), исследования устойчивости 623—627, 676
Гейссер (Geisser S.), фидуциальные выводы 215
Гентер (Guenther W. C.), несмещенные критерии для дисперсии нормального распределения 286
Гёрлэнд (Gurland J.), оценивание в распределениях Неймана типа A и отрицательно-биномиальном (упражнение 18.28) 107; оценивание функциональных и структурных зависимостей 511, 543
Гест (Guest P. G.), ортогональные полиномы 480
Гёфдинг (Hoeffding W.), эффективность критериев ОП и X^2 в мультиномиальных выборках 370, 588; нормальное распределение и распределение перестановок 636; совместное распределение ранговых корреляций 644; критерий независимости 647; оптимальные свойства критериев, использующих нормальные метки 652, 668; асимптотическое распределение математических ожиданий порядковых статистик 652; распределение критерия Вилкоксона 663
Гехан (Gehan E. A.), критерий Вилкоксона для цензурированных выборок 707
Гиббонс (Gibbons J. D.), отсутствие нормальности и критерий знаков 690
Гибсон (Gibson W. M.), оценивание функциональной зависимости 539
Гилби (Gilby W. H.), данные об обеспеченности и интеллекте (пример 33.7) 747
Гилфорд (Guilford J. P.), тетрахорическая корреляция 410
Гипергеометрическое распределение (hypergeometric distribution), обратный выбор 795
Гипотезы (hypotheses) статистические 219, 245—246; параметрические и непараметрические 220; простые и сложные 221; степени свободы, ограничения 221; критические области и альтернативные гипотезы 221—223; нулевая гипотеза 222, сноска; см. Критерии для проверки гипотез
Гири (Geary R. C.), «близость» оценки 22; асимптотическая минимизация обобщенной дисперсии МП-оценками 86; функциональные и структурные зависимости 526, 546, 550, (упражнения 29.2, 29.15) 554, 556; проверка нормальности 617; устойчивость 622—625
Гиршик (Girshick M. A.), решающие функции 827; последовательный анализ (упражнения 34.8, 34.9, 34.20) 829, 831
«Гладкие» критерии согласия («smooth» tests of fit) 593—602
Гласс (Glass D. V.), критерий Дёрбина о диагональной пропорциональности в таблицах $r \times r$ (упражнение 33.25) 788
Гнанадесикан (Gnanadesikan M.), логистические оценки (упражнение 18.29) 107; цензурированные гамма-выборки 706
Говиндараджулу (Govindarajulu Z.), мощность критерия Вилкоксона 667; библиография по испытаниям на долговечность 705; цензурированное двойное экспоненциальное распределение 706
Годэмб (Godambe V. P.), обобщение МГД (упражнение 18.32) 108
Голдбергер (Goldberger A. S.), двухэтапный метод НК (упражнение 28.24) 499
Голдмэн (Goldman A. S.), двойной выбор 826
Гохеен (Goheen H. W.), тетрахорическая корреляция 410
Гоэн (Goen R. L.), урезанное пуассоновское распределение 706, (упражнение 32.19) 716
Граббс (Grubbs F. F.), критерии для выпадающих наблюдений 710
Грама-Шарлье ряды (Gram-Charlier series), типа A, эффективность метода моментов (упражнение 18.28) 107
Гранди (Grundy P. M.), группированные урезанные и цензурированные нормальные данные 703
Грейбилл (Graybill F. A.), квадратичные формы от нецентральных нормальных величин 308; двойной выбор для оценивания нормальной дисперсии 826
Гринберг (Greenberg B. G.), порядковые статистики 687; цензурированные выборки 704, 706, (упражнение 32.25) 718
Гринвуд (Grenwood M.), данные о прививке (пример 33.1) 722
Гринхауз (Greenhouse S. W.), таблицы стьюдентизованного максимального абсолютного отклонения 710
Группировка (grouping), поправки для МП-оценок (упражнения 18.24—18.25) 105—106; для инструментальных переменных 532—540, 549; урезанные и цензурированные данные 703—705
Гудмэн (Goodman L. A.), толерантные интервалы 698; связи в категоризованных данных 730, 752, 757—758, (упражнение 33.11) 784; таблицы 2×2 744; взаимодействия в таблицах с несколькими входами 781
Гупта (Gupta S. S.), оценки параметров логистического распределения (упражнение 18.29) 107; логистические порядковые статистики (упражнение 32.12) 714
Гхош (Ghosh J. K.), полнота 255; инвариантность и достаточность 344, 819

Дадмэн (Dudman J.), логистические порядковые статистики (упражнение 32.12) 714
Даниэлс (Daniels H. E.), асимптотическая нормальность и эффективность МП-оценок 69; неединственность МП-оценок (упражнение 18.33) 108; минимизация обобщенной дисперсии НК-оценками 118; оценивание функции от нормального коэффициента корреляции 393; коэффициенты корреляции и беспорядок 640, 754; совместное распределение ранговых корреляций 644, (упражнение 31.6) 683; теорема о корреляциях (упражнение 31.3) 683
Данциг (Dantzig G. B.), отсутствие критерия для нормального среднего с мощностью, не зависящей от неизвестной дисперсии 265
Дар (Dar S. N.), отношение нецентральных F-величин 341
Дарлинг (Darling D. A.), критерии согласия 605, 616; распределение отношения между суммой и экстремальным значением 710—711
Дармуа (Darmois G.), достаточность 46, 48
Даунтон (Downton F.), порядковое оценивание НК 126; распределение экстремальных значений 706
Двойная выборка (double sampling) 810, 824—826, (упражнения 34.15—34.17) 830—831

Двойное экспоненциальное (Лапласа) распределение (double exponential (Laplace) distribution), МП-оценка среднего (пример 18.7) 70; как альтернатива при проверке нормальности (пример 22.5) 229; критерий знаков, асимптотически наиболее мощный для параметра сдвига (упражнение 32.5) 713
Двумерное нормальное распределение (bivariate normal distribution), МП-оценка коэффициента корреляции ϱ (пример 18.3) 62; неопределенность МП-оценки для функции от ϱ (пример 18.4) 66; асимптотическая дисперсия для $\hat{\rho}$ (пример 18.6) 70; МП-оценки различных комбинаций параметров (пример 18.14) 87; (пример 18.15) 89; (упражнения 18.11—18.14) 103; графики доверительных интервалов для ϱ 164; доверительные интервалы для отношения дисперсий (упражнение 20.19) 182; мощности критериев для ϱ (пример 22.7) 233; совместная х. ф. квадратов случайных величин (пример 26.1) 379; линейная регрессия для квадратов случайных величин (пример 26.3) 381, (пример 26.5) 382; оценивание коэффициента ϱ 392—394; доверительные интервалы и критерии для ϱ 394—395; критерии независимости и регрессионные критерии 395—396; корреляционные отношения (пример 26.8) 398; регрессии линейные, гомоскедастические (упражнение 26.2) 418, (пример 28.1) 465; оценивание ϱ, когда исходное распределение дихотомизировано (упражнения 26.3—26.4) 418; распределение коэффициента регрессии (упражнение 26.9) 419; МП-оценка в бисериальном случае (упражнения 26.10—26.12) 419—420; критерий ОП для ϱ (упражнение 26.15) 421; МП-оценка и критерии ОП для коэффициента корреляции, общего для двух распределений (упражнения 26.19—26.22) 421—422; совместное распределение суммы квадратов (пример 28.2) 466; характеризации 469—471, (упражнения 28.7—28.11) 494; устойчивость критериев для ϱ 627; эффективность критериев независимости 633—635, 644—647; эффективность критериев для регрессии 649—653; РНМН-критерии для гипотезы $\varrho=0$ (упражнение 31.21) 686; интерпретация коэффициента сопряженности 746—747; преобразования и канонические корреляции 761—762
Двухвыборочные критерии (two-sample tests) 653—675
Де Гроот (De Groot M. H.), последовательная выборка по качественному признаку 819 (сноска)
Дейли (Daly J. F.), наименьшая доверительная область 175; несмещенные критерии ОП для независимости 329
Демпстер (Dempster A. P.), фидуциальные выводы 211, 212
Ден Бродер (Den Broeder G. G., Jr.), урезанные и цензурированные гамма-величины 706
Дёрбин (Durbin J.), обобщение МГД (упражнение 18.32) 108; геометрия теории НК 116; дополнительная информация при регрессии (упражнения 28.17, 28.18) 496—497; инструментальные переменные 531, 543, (упражнение 29.17) 557; критерии согласия и распределения равномерных порядковых статистик (упражнение 30.17) 621; устранение мешающего параметра путем отбрасывания достаточной статистики (упражнение 30.18) 621; критерий диагональной пропорциональности в таблице $r\times r$ (упражнение 33.25) 788
Джейкобсон (Jacobson J. E.), таблицы критерия Вилкоксона 662
Джеймс (James G. S.), таблицы для задачи о двух средних 202
Дженкинс (Jenkins W. L.), номограмма для тетрахорической корреляции 410
Джеффрейс (Jeffreys, Sir Harold), байесовские интервалы 205, 210
Джогдео (Jogdeo K.), критерии, использующие случайные нормальные отклонения 653
Джокемс (Jochems D. B.), двухэтапный метод НК (упражнение 28.24) 499
Джонкхир (Jonckheere A. R.), критерий для k-выборок 677—678
Джонс (Johns M. V., Jr.), эффективное оценивание с помощью порядковых статистик 703
Джонсон (Johnson N. L.), «точные» оценки 22; нецентральное t-распределение 342; нецентральное χ^2 и пуассоновское распределения (упражнение 24.19) 350; интегральное вероятностное преобразование с оцененными параметрами 593, (упражнение 30.10) 620; библиография по порядковым статистикам 687; ПКОВ 806, (упражнение 34.5) 828; обзор последовательного анализа 827
Джоуэтт (Jowett G. H.), оценивание функциональной зависимости 539
Джэксон (Jackson J. E.), библиография по последовательному анализу 827
Диксон (Dixon W. J.), эффективность критериев 355; мощность критерия Вилкоксона 667; критерий знаков 689—692; цензурированные нормальные выборки 704; выпадающие наблюдения 708—710
Димер (Deemer W. L., Jr), урезание и цензурирование для экспоненциального распределения 705, (упражнение 32.16) 715
Дискретность (discontinuity), и доверительные интервалы 146—148, 164—165; критерии 226; влияние на непараметрические критерии 681—683; поправка в таблицах 2×2 744—745
Дисперсионные критерии (dispersion tests), биномиальный и пуассоновский 774—776
Дисперсия (variance), нижняя граница для нее, см. Минимальная граница дисперсии; обобщенная, см. Обобщенная дисперсия
Доверительная полоса (confidence belt) 141; см. Доверительные интервалы
Доверительное распределение (confidence distribution) 144
Доверительные границы (confidence limits) 141; см. Доверительные интервалы
Доверительные интервалы (confidence intervals), общее 139—182; распределения с одним параметром 139—165; границы 141; полоса 141; коэффициент 141; графическое представление 144, 147, 165, 167, 171; центральные и нецентральные 145; дискретные распределения 146—148, 164—165; консервативные 146—148; для больших выборок 148—151, 155; вложенность 150—151; трудные случаи 151—155, (пример 20.7) 172—174, (упражнение 28.21) 497; кратчайшие 158—162; минимальная средняя длина для больших выборок 158—161; наиболее селективные 161, 170, 277; д. и. и критерии 161—162, 277; несмещенные д. и. 162, 277; таблицы и гра-

фики 162—164; распределения с несколькими параметрами 165—169; выбор статистики 169; стьюдентизованные д. и. 170—174; совместные интервалы для нескольких параметров 174—175; д. и., когда область изменения зависит от параметра (упражнения 20.13—20.16) 180—181; задача о двух средних 190—203, (упражнение 21.9) 218; критическое обсуждение 207—211, 216; д. и. для непрерывной ф. р. 611—613; д. и. для параметра сдвига (непараметрические) 658, (упражнения 31.24—31.25) 686; д. и. для квантилей 692—694
Доверительные области (confidence regions) 174—175, (упражнение 20.5) 179; для регрессии 487—493
Доверия коэффициент (confidence coefficient) 141; см. Доверительные интервалы
Додж (Dodge H. F.), двойная выборка 810
Доксум (Doksum K. A.), критерии, использующие случайные нормальные отклонения 653
Дополнительная информация (supplementary information), в регрессии (упражнения 28.16—28.18) 495—497; инструментальные переменные 530—543
Дорфф (Dorff M.), оценивание функциональных и структурных зависимостей 511, 543
Достаточность (sufficiency), общее 40—53; определение 40—41; критерий факторизации 41; д. и МГД 42—43; функциональная зависимость между достаточными статистиками 44; д. и МД-оценка 44; распределения, обладающие достаточными статистиками 45—47; при нескольких параметрах 47—49; когда область изменения варианты зависит от параметра 49—52, (упражнение 24.17) 349; одномерная и совместная 47, (упражнения 17.13, 17.25) 55, 57; распределение достаточных статистик в выборках из экспоненциального семейства распределений (упражнение 17.14) 55; д. и МП-оценка 59—61, (пример 18.5) 67, 80—82; д. и НКО критериев 230, 238—241, (упражнения 22.11, 22.13) 248; оптимальность критерия, основанного на достаточных статистиках 250—251; д. и подобные области 254, 263; минимальная 259—263, (упражнение 18.13) 103, (упражнение 23.31) 300; вспомогательные статистики и квазидостаточность 292; независимость и полнота (упражнения 23.6—23.7) 296; д. и критерии ОП 329; д. и инвариантность 344; д. и мешающие параметры (упражнение 30.18) 621; д. и урезание 699—700
Достаточные статистики (sufficient statistics) см. Достаточность
Дункан (Duncan A. J.), графики функции мощности для линейной гипотезы 340
Дынкин Е. Б., минимальная достаточность 261
Дэвид (David H. A.), таблицы стьюдентизованного отклонения и стьюдентизованного интервала 709—710
Дэвид С. (David S. T.), трудности в теории доверительных интервалов 152; совместное распределение ранговых корреляций 644
Дэвид Ф. (David F. N.), модифицированные МХК-оценки (упражнение 19.14) 138; графики для коэффициента корреляции двумерного нормального распределения 164, 394; смещенность критерия для проверки гипотезы о коэффициенте корреляции нормального распределения 394; критерий серий в дополнение к критерию согласия X^2 593, (упражнения 30.7, 30.9) 619—620; вероятностное интегральное преобразование с оцененными параметрами 591, (упражнение 30.10) 620; «гладкие» критерии согласия 596; библиография о порядковых статистиках 687
Дэвис (Davis R. C.), достаточность, когда область изменения зависит от параметра 51
Дэрроч (Darroch J. N.), взаимодействия в таблицах с несколькими входами 781

Женщин физические данные (women's measurements and sizes), сведения (таблицы 26.1—26.2) 376—377

Зависимость и взаимозависимость (dependence and interdependence) 373—375
Задача о двух средних (two means problem) см. Нормальное распределение
Закриссон (Zackrisson U.), t-распределение Стьюдента в составных нормальных выборках 625
Залокар (Zalokar J.), таблицы стьюдентизованного максимального абсолютного отклонения 710
Зискинд (Zyskind G.), двухэтапный метод НК (упражнение 28.24) 499
Значимости критерии (significance tests), см. Критерии для проверки гипотез, Критерии согласия
Значимости уровень (significance level) 222, сноска

Ивашкевич (Iwaskiewicz K.), таблицы функции мощности t-критерия Стьюдента 342
Идентифицируемость (identifiability) (пример 18.4) 67; в структурных зависимостях 506—508, 523, 526
Инвариантные критерии (invariant tests) 325, 344—345; и ранговые критерии 637—638, 647, 659
Инверсии (inversions of order) 640
Инструментальные переменные (instrumental variables) 530—543
Интервальное оценивание (interval estimation) 139—218
Информация (information), количество 25; матрица 49
Ирвин (Irwin J. O.), урезанное пуассоновское распределение (упражнения 32.23—32.24) 717; выпадающие наблюдения 708; компоненты X^2 773
Испытания на долговечность (life testing) 705
Ист (East D. A.), таблицы доверительных интервалов для нормальной дисперсии 163
Исчерпывающий (exhaustive) 260, сноска
Иэйтс (Yates F.), фидуциальные выводы 211; проверка гипотез о различии между регрессиями (упражнение 28.14) 495; критерии, использующие нормальные метки 652, 673; данные о зубах с неправильным прикусом (пример 33.5) 739; поправка на непрерывность в таблицах 2×2 745; использование меток в таблицах $r \times c$ 760, (упражнение 33.12) 784

Каврук (Kavruck S.), тетрахорическая корреляция 410
Канонический анализ таблиц $r \times c$ (canonical analysis of $r \times c$ tables) 761—769
Кастен (Kasten E. L.), графики для критерия независимости в таблицах 2×2 741
Кастенбаум (Kastenbaum M. A.), гипотезы в таблицах с несколькими входами 780
Категоризованные данные (categorized data) 719—790; меры связи 719—725, 745—749, 752, (упражнения 33.1—33.4, 33.6) 781—782; частная связь 725—730; вероятностные интерпретации 730—733; асимптотические критерии независимости 733—736, 749—751; точный критерий независимости, различные модели 736—744, 749; поправка на непрерывность 744—745, (упражнение 33.5) 782; упорядоченные таблицы, ранговые меры связи 752—758, (упражнения 33.10—33.11) 784; упорядоченные таблицы, методы, использующие метки и канонический анализ 758—769, (упражнения 33.12—33.14) 784—785; применение метода НК 760; разбиение величины X^2, канонические компоненты 769—774, (упражнение 33.13) 785; биномиальный и пуассоновский дисперсионные критерии 774—776; таблицы с несколькими входами 776—781; критерий полной симметрии (упражнение 33.25) 788; критерий для проверки однородности двух маргинальных распределений (упражнение 33.24) 788; МП-оценка групповых вероятностей (упражнение 33.29) 789; критерий для проверки однородности маргинальных распределений в таблице 2^C (упражнение 33.30) 790
Катц (Katz L.), мощность критерия в таблицах 2×2 743
Катц (Katz M., Jr.), минимальная достаточность 260, сноска
Кац (Kac M.), сравнение критерия согласия X^2 и критерия Колмогорова 615; критерии нормальности 618
Качественные признаки (attributes), выборочное изучение (sampling of), см. Биномиальное распределение
Квадратичные формы (quadratic forms), среднее и дисперсия (упражнения 19.3, 19.9) 136—137; нецентральное χ^2-распределение квадратичной формы 308, (упражнение 24.10) 347
Квантили (quantiles), критерии и доверительные интервалы для них 687—695
Кейб (Kabe D. G.), уточнение регрессионного анализа 474
Кейл (Kale B. K.), итерационное вычисление МП-оценки 76, 91
Кейнес (Keynes J. M.), характеризация распределений формой МП-оценок (упражнения 18.2—18.3) 100
Кейпон (Capon J.), критерии различия в масштабе 674; локально наиболее мощные ранговые критерии (упражнение 31.17) 685
Кемперман (Kemperman J. H. B.), толерантные области 698
Кемпторн (Kempthorne O.), таблицы Тэнга 340
Кендалл (Kendall M. G), геометрическая интерпретация теории НК 116; оценивание функции от нормального коэффициента корреляции 393; линейная регрессия с однородными ошибками наблюдений 552—553, (упражнение 29.13) 556; распределения ранговых корреляций 639—644; поправки для ранговых коэффициентов корреляции 683; ранговая мера связи 755—757
Кимбелл (Kimball A. W.), вычисление разбиения величины X^2 773
Кимбелл Б. (Kimball B. F.), МП-оценка в распределении экстремальных значений (упражнение 18.6) 101
Кифер (Kiefer J.), границы для дисперсий оценок 33; пример, когда МП-оценка не существует (упражнение 18.34) 108; состоятельность МП-оценок 514, 517; сравнение критерия согласия X^2 и критерия Колмогорова 615; критерии нормальности 618; распределение критерия независимости Гёфдинга 647
Кларк Р. (Clark R. F.), таблицы доверительных границ для биномиального параметра 163
Кларк (Clark V. A.), логистические порядковые статистики (упражнение 32.12) 714
Клетт (Klett G. W.), таблицы доверительных интервалов для нормальной дисперсии 163; кратчайшие доверительные интервалы для нормальной дисперсии (упражнение 20.10) 180
Клитическая кривая (clitic curve) 462
Клоппер (Clopper C. J.), графики доверительных интервалов для биномиального параметра 162
Клоц (Klotz J.), распределение критерия, основанного на нормальных метках 668; критерии c_1 и t 673; критерий для различия в масштабе 674; критерии симметрии 681
Клюнье-Росс (Clunies-Ross C. W.), двухэтапный метод НК (упражнение 28.24) 499
Ковариация (covariance) 378; см. Корреляция, Регрессия
Кокрэн (Cochran W. G.), критерий знаков (упражнение 25.1) 371, 687—689; уточнение регрессионного анализа 474; критическая область для критерия X^2 562; выбор классов для критерия X^2 588; предельная функция мощности критерия X^2 (упражнение 30.4) 618; выпадающие наблюдения 710; X^2 в таблицах 2×2 745; разбиение X^2 773; дисперсионные критерии 775; критерий для частной связи 781; критерий для проверки однородности маргинальных распределений в таблицах 2^C (упражнение 33.30) 790
Кокс (Cox C. P.), ортогональные полиномы 480
Кокс Д. (Cox D. R.), оценивание линейной связи между функциями от параметров независимых биномиальных распределений (упражнение 17.25) 57; доверительное распределение 144; условные критерии 294; статистики ОП 331; АОЭ и максимальная потеря мощности 366; регрессия эффективной оценки относительно неэффективной (упражнение 25.7) 371; экспоненциальные метки 672; АОЭ простых критериев случайности (упражнения 31.10—31.12) 684; последовательные процедуры для сложной гипотезы H_0 819, (упражнение 34.21) 832; двойная выборка 826, (упражнения 34.15—34.17) 830—831
Количество информации (amount of information) 25
Коллигации коэффициент (colligation coefficient) 723
Колмогоров А. Н., критерий согласия 605—617
Колоджейчик (Kolodziejczyk S.), общая линейная гипотеза 334; таблицы функции мощности t-критерия Стьюдента 342

Комбинация критериев (combination of tests) (упражнение 30.9) 620
Компоненты X^2 (components of X^2) 601—602; в таблицах $r \times c$ 769—774
Кондо (Kondo T.), стандартная ошибка критерия X^2 в таблице $r \times c$ 750
Конейн (Konijn H. S.), АОЭ критериев 358; линейные преобразования независимых величин 646
Коннелл (Connell T. L.), двойная выборка для нормальной дисперсии 826
Контролируемые переменные (controlled variables) 544—545
Корнфилд (Cornfield J.), фидуциальные выводы 215; таблицы стьюдентизованного максимального абсолютного отклонения 710
Корреляция (correlation), между оценками 35—36; общее 373—422; к. и взаимозависимость 373—374, 385; к. и причинность 374—375; коэффициент к. 384—385; вычисление коэффициента к. (примеры 26.6—26.7) 386—389; диаграмма разброса 389; стандартные ошибки 390; оценивание и проверка гипотез в нормальных выборках 392—395; корреляционные отношения 396; к. и линейность регрессии 396—398, (упражнение 26.24) 422; вычисление корреляционного отношения (пример 26.9) 399—400; критерии ОП для коэффициента к., корреляционного отношения и линейности регрессии 400—403; внутриклассовая 403—407; тетрахорическая 407—411; бисериальная 411—416; точечно-бисериальная 416—417; возрастание коэффициента к. при поправках Шеппарда на группировку (упражнение 26.16) 421; изменение к. вследствие погрешностей измерений (упражнение 26.17) 421; «ложная» (упражнение 26.18) 421; матрица 424; определитель 424; см. также Множественная корреляция, Частная корреляция и регрессия, Регрессия
Корреляция внутриклассовая (intra-class correlation), 403—407, (упражнение 26.14) 420—421; в многомерном нормальном распределении (упражнение 27.17) 459
Коши распределение (Cauchy distribution), непригодность выборочного среднего в качестве оценки медианы 15; состоятельность выборочной медианы (пример 17.5) 22; МГД для параметра сдвига (пример 17.7) 26; МП-оценки и оценки, основанные на порядковых статистиках, для параметра сдвига (пример 18.9) 76—77; проверка простой гипотезы H_0 о медиане (пример 22.4) 228—229, (упражнение 22.4) 248; полнота 255, (упражнение 23.5) 295
Коэн (Cohen A. C., Jr.), урезание и цензурирование в нормальных распределениях 703; урезание и цензурирование в пуассоновских распределениях 706, (упражнение 32.22) 717; урезанное гамма-распределение 706
Крамер (Gramér H.), МГД 24; эффективность и асимптотическая нормальность МП-оценок 68; распределение критерия согласия X^2 567; критерий согласия 603; коэффициент связи 747
Крамер К. (Kramer K. H.), доверительные интервалы для многомерной нормальной множественной корреляции 457
Крамера — Рао неравенство (Cramér — Rao inequality), см. Минимальная граница дисперсии
Крауз (Crouse C. F.), критерий Муда 674
Кривые кусочно-линейные (segmented curves), оценивание по методу НК 474
Кризи (Creasy M. A.), доверительные интервалы для линейной функциональной зависимости 519
Критерии для проверки гипотез (tests of hypotheses) 219—372; критерии и доверительные интервалы 161—162, 277; размер 222, 245—247; мощность 223—224, 242—243; НКО 225; простая гипотеза H_0 против простой H_1 225; рандомизация в дискретном случае 226; НКО и достаточные статистики 230—231, 238—241; мощность и эффективность оценивания 231—233, (упражнение 22.7) 248; простая гипотеза H_0 против сложной H_1 233; РНМ критерии 234; отсутствие РНМ критерия против двусторонней H_1 234—235; РНМ критерии в случае нескольких параметров 235—241, (упражнение 22.11) 248; одно- и двусторонние критерии 243—245; оптимальное свойство достаточных статистик 250—251; подобные области и критерии 251—254, 263—269, 275—276; существование подобных областей 252—254, (упражнения 23.1—23.3) 295, (упражнение 23.29) 300; подобные области, достаточность и ограниченная полнота 254—255, 263—269; наиболее мощные подобные области 263; неподобные критерии для сложной гипотезы H_0 269; смещение 270; несмещенные критерии и подобные критерии 275—276; старая терминология 277; РНМН критерии для экспоненциального семейства 278—292; вспомогательные статистики и условные критерии 292—295; критерии ОП 300—331; несмещенные инвариантные критерии для параметров сдвига и масштаба 324—329, (упражнение 24.15) 349; общая линейная гипотеза 333—344; последовательные критерии 798—819, 824—827; см. также Асимптотическая относительная эффективность, Гипотезы, Критерии отношения правдоподобия
Критерии согласия (tests of fit) 558—621, критерии ОП и Пирсона для простой гипотезы H_0 558—563; обозначение X^2 561 (сноска); сложная гипотеза H_0 563—567; выбор классов для критерия X^2 574—578; 583—589; моменты статистики X^2 578—579, (упражнения 30.3, 30.5) 618—619; состоятельность и несмещенность критерия X^2 579—581; предельная мощность критерия X^2 581—583; использование знаков отклонений 590—591, (упражнения 30.7—30.9) 619—620; критерии, отличные от X^2 591—617; «гладкие» критерии 593—595; связь между X^2 и «гладкими» критериями 600; компоненты X^2 601—602; критерии, основанные на выборочной ф. р. 602—617; критерий Смирнова 602—605; критерий Колмогорова 605—611; сравнение критерия X^2 и критерия Колмогорова 613—616; критерии нормальности 617—618
Критерий знаков (sign test), АОЭ (упражнения 25.1—25.2) 370—371, 687—692; последовательный 826
Критерий отношения правдоподобия (ОП) (likelihood ratio (LR) test), нормальное распределение (пример 22.8) 236; общее 301—331; к. ОП и МП-оценки 301—302, 323—324; возможность отсутствия подобия 302—304; аппроксимации распределений 305, 308—314; асимптотическое распределение 308—310, асимптотическая мощность и таблицы 310—313, 339—340;

к. ОП, когда область изменения зависит от параметра 317—322; (упражнения 24.8—24.9) 347; свойства 322—330; состоятельность 310—311, 322—323; несмещенность 323—324. 329—330, (упражнения 24.14—24.18) 349—350; другие свойства 329—330; бесполезный критерий (пример 24.7) 330; к. ОП для линейных гипотез 334—336; функция мощности 339—341; оптимальные свойства 342—344; эффективность в мультиномиальных выборках 370; к. ОП для нормальной корреляции и регрессии 400—403, 450—451; критерий согласия 559—563, (упражнение 30.11) 620; независимость в таблицах 2×2 733—736; в таблицах $r \times c$ 752

Критическая область (critical region), см. Критерии для проверки гипотез

Кришнан (Krishnan M.), несмещенные критерии типа M 272; дважды нецентральная t-величина 342

Кроу (Crow E. L.), таблицы доверительных границ для доли успехов и для пуассоновских параметров 163

Крускал (Kruskal W. H.), геометрия метода НК 117; линейные регрессии, коэффициенты корреляции и корреляционные отношения (упражнение 26.24) 422; история ранговой корреляции 640; критерий Вилкоксона 660—662, (упражнение 31.15) 685; k-выборочные критерии 676; совпадения 682; связь в категоризованных данных 730, 752, 757—758, (упражнение 33.11) 784

Крэйг (Craig A. T.), полнота достаточных статистик 257; полнота и независимость (упражнения 23.8—23.9) 296; критерий ОП для равномерного распределения (упражнение 24.9) 347; достаточность в экспоненциальном семействе, когда область изменения зависит от параметров (упражнение 24.17) 349

Кудо (Kudo A.), отклонение выпадающих наблюдений 709

Куисенберри (Quesenberry C. P.), критерии для выпадающих наблюдений 709

Куллдорф (Kulldorff G.), цензурированные экспоненциальные выборки 706

Купмэн (Koopman B. O.), распределения с достаточными статистиками 46, 48

Куртическая кривая (kurtic curve) 462

Кэнуй (Quenouille M. H.), поправки на смещение 19—21, (упражнения 17.17—17.18) 55; фидуциальные парадоксы 211; случайные экспоненциальные отклонения (пример 30.2) 576; разбиение X^2 774

Кэтти (Katti S. K.), оценивание в распределениях Неймана типа A и отрицательно-биномиальном (упражнение 18.28) 107

Лайонс (Lyons T. C.), тетрахорическая корреляция 410

Лапласа преобразование (Laplace transform) 256 (сноска)

Лаха (Laha R. G.), линейность регрессии 554

Лача (Latscha R.), таблицы точного критерия независимости в таблицах 2×2 741, 743

Левонтин (Lewontin R. C.), МП-оценка числа классов в мультиномиальном распределении (упражнение 18.10) 102

Лежандра полиномы (Legendre polynomials) 594 (сноска)

Ле Кам (Le Cam L.), сверхэффективность 69

Леман (Lehmann E. L.), пример бессмысленной несмещенной оценки 19, (упражнение 17.26) 57; проверка гипотез 219; оптимальное свойство критерия, основанного на достаточных статистиках 250; полнота 254; полнота линейного экспоненциального семейства 256; минимальная достаточность 260—261; полнота и подобные области 263; задача о двух средних (пример 23.10) 268; неподобные критерии для сложной гипотезы H_0 269; несмещенные критерии 277; РНМН критерии для экспоненциального семейства 278; полнота семейства Коши (упражнение 23.5) 295; минимальная достаточность для биномиального и равномерного распределений (упражнения 23.11—23.12) 297; РНМН критерий для нормального распределения (упражнение 23.16) 297; РНМН критерии для распределений Пуассона и биномиального (упражнения 23.21—23.22) 298; РНМ критерии для экспоненциального и равномерного распределений (упражнения 23.25—23.26) 299; бесполезность критерия ОП (пример 24.7) 330; оптимальные свойства критериев ОП 343; асимптотически РНМ критерии для нормального среднего (пример 25.1) 353; эффективность критериев 355; РНМ инвариантные критерии для коэффициентов корреляции и множественной корреляции 394, 457; доверительные интервалы для регрессии (упражнение 28.21) 497; распределение критерия согласия X^2 570, 573; полнота порядковых статистик 632 (сноска); несмещенность ранговых критериев независимости 647; ранговые критерии 653; распределение критерия Вилкоксона 663; доверительные интервалы, основанные на критериях Вилкоксона 663 (упражнения 31.24—31.25) 686; состоятельность критерия Вилкоксона 663; АОЭ критерия Вилкоксона 667; мощность критерия Вилкоксона 667; устойчивое оценивание 673; несмещенные двухвыборочные критерии 674; критерии симметрии 679; АОЭ критерия Вилкоксона и критерия, основанного на нормальных метках (упражнение 31.23) 686; оптимальные свойства критерия знаков 689; минимальная АОЭ критерия знаков (упражнение 32.1) 712; критерий знаков как асимптотически наиболее мощный критерий сдвига для двойного экспоненциального распределения (упражнение 32.5) 713; последовательная полнота 819 (сноска)

Лемер (Lehmer E.), таблицы функции мощности для линейной гипотезы 340

Либерман (Lieberman G. F.), таблицы нецентрального t-распределения 342

Линдеберг (Lindeberg F. W.), ранние работы о ранговой корреляции 640

Линдли (Lindley D. V.), поправки на группировку в МП-оценках (упражнения 18.24—18.25) 105—106; таблицы доверительных интервалов для нормальной дисперсии 163; согласование фидуциального вывода с байесовским 212—215, (упражнение 21.11) 218; выбор размера критерия 247; условные критерии 294; несостоятельность МП-оценок в задаче о функциональной зависимости 514, 516; контролируемые переменные 544; ошибки наблюдений и линейность регрессии 551, 554 (упражнения 29.12—29.14) 556

Линейная модель (linear model) 111—126; критерии, требующие предположения о нормальности 120—122; более общий случай 126, (упражнения 19.2—19.3, 19.5, 19.9) 136—137; общая линейная гипотеза 332; каноническая форма 333; статистика ОП в л. м. 334; л. м. и теория НК 336; функция мощности критерия ОП 339—341; оптимальные свойства критерия ОП 342—344; л. м. и регрессия 472—493; смысл термина «линейная» 473—474; доверительные интервалы и критерии для параметров 482—487, (упражнения 28.15—28.22) 495—498; доверительные области для линии регрессии 487—493; дополнительная информация для регрессии (упражнения 28.16—28.18) 495—497; уточнение при одном дополнительном наблюдении (упражнение 28.19) 497; см. Наименьших квадратов оценки, Регрессия
Линейная регрессия (linear regression), см. Регрессия
Линник Ю. В., задача о двух средних 269
Липпс (Lipps G. F.), ранние работы о ранговой корреляции 640
Ллойд (Lloyd E. H.), порядковые НК-оценки параметров масштаба и сдвига 126, (упражнение 19.10) 137
Логистическое распределение (logistic distribution), МГД (упражнение 17.5) 54; МП-оценка (упражнение 18.29) 107; АОЭ критерия Вилкоксона (упражнение 31.17) 685; х. ф. и семиинварианты порядковых статистик (упражнение 32.12) 714
Логнормальное распределение (lognormal distribution), МП-оценка (упражнения 18.7—18.9, 18.19—18.20, 18.23) 101—102, 104—105
Локс (Locks M. O.), таблицы нецентрального t-распределения 342
Лоли (Lawly D. N.), аппроксимации распределения статистики ОП 313
Лорент (Laurent A. G.), цензурированные экспоненциальные выборки 706
Лотон (Lawton W. H.), задача о двух средних 201
Льюис (Lewis B. N.), таблицы с несколькими входами 777
Льюис П. (Lewis P. A. W.), таблицы критерия согласия 605
Льюис Т. (Lewis T. O.), НК-оценки в случае вырождения 124
Лэнкастер (Lancaster H. O.), поправки на непрерывность при сложении значений величины X^2 745; полихорическое оценивание 752; канонические корреляции и преобразование в двумерное нормальное распределение 761—762; компоненты X^2 769, (пример 33.13) 771—772, (упражнение 33.15—33.18) 786; таблицы с несколькими входами (пример 33.14) 778, 781, (упражнение 33.8) 782—783, (упражнения 33.27—33.28) 788—789

Мадански (Madansky A.), кратчайшие доверительные интервалы 162; функциональные и структурные зависимости 511, 528, 539, (упражнение 29.16) 556; толерантные интервалы 628; таблицы с несколькими входами (упражнения 33.24, 33.30) 788, 790
Майерс (Myers M. H.), последовательные t-критерии 819
Майтра (Maitra A. P.), достаточность 48
Мак Кей (McKay A. T.), распределение экстремальных отклонений от среднего (упражнение 32.21) 716—717
Мак Кендрик (McKendrick A. G.), оценивание в цензурированных выборках (упражнение 32.24) 717
Маккиннон (Mackinnon W. J.), таблицы критерия знаков и доверительных интервалов для медианы 689, 695
Мак Корнак (McCornack R. L.), таблицы критерия симметрии Вилкоксона 681
Максимального правдоподобия (МП-) оценки (maximum likelihood (ML) estimators), общее 58—108; принцип МП 58; МП-оценки и МГД-оценки 59—61, (упражнение 18.32) 108; МП-оценки и достаточность (пример 18.5) 59—61; единственность при наличии достаточных статистик 59—61, 80—82; отсутствие несмещенности (пример 18.1) 61—62, (пример 18.11) 82; асимптотические оптимальные свойства 62; состоятельность и несостоятельность 64—67, 83—84, (пример 18.16) 92—93, (упражнение 18.35) 108; 514, 516—517; случаи неопределенности (пример 18.4) 66—67, (упражнения 18.17, 18.23, 18.33—18.35) 103—105, 108; эффективность и асимптотическая нормальность 67—72, 83—86; асимптотическая дисперсия, равная МГД 68—70; упрощение асимптотической дисперсии и дисперсионной матрицы при наличии достаточности 71, 84; семиинварианты 72—75; последовательные аппроксимации 75—79; случай нескольких параметров 70—91; асимптотическая минимизация обобщенной дисперсии 85—86; неодинаковые исходные распределения 91—93; МП-оценки параметров сдвига и масштаба 93—97; характеризация исходных распределений, имеющих МП-оценки заданной формы (упражнения 18.2—18.3) 100; МП-оценки для целочисленных параметров (упражнения 18.21—18.22) 104—105; поправки на группировку (упражнения 18.24—18.25) 105—106; совпадение с НК-оценками в нормальном случае 109, 116; случай структурной и функциональной зависимости 506—511, 512—518; применение для проверки согласия 570—574; для задач урезания и цензурирования 701—702; в последовательных критериях (упражнение 34.21) 831—832
Мак Стьюарт (Mac Stewart W.), мощность критерия знаков 690
Манн (Mann H. B.), таблицы Тэнга 340; выбор классов для критерия согласия X^2 576, 583—587; несмещенность критерия X^2 580; дисперсия величины X^2 (упражнение 30.3) 618; критерий ранговой корреляции для тренда 647—649, (упражнение 31.8) 684; критерий Вилкоксона 660
Мантел (Mantel N.), доверительные интервалы для нелинейных функций 175
Мариц (Maritz J. S.), бисериальная корреляция 416
Марсалья (Marsaglia G.), квадратичные формы от нецентральных нормальных величин 308; многомерная нормальная регрессия (упражнение 27.21) 460
Масштабные параметры (scale parameters), см. Сдвига и масштаба параметры
МГД, см. Минимальная граница дисперсии
МД, см. Минимальной дисперсии оценки
Медиана (median), критерий знаков для нее 689—692, (упражнение 32.1) 712; доверительные интервалы для нее 694—695, (упражнение 32.4) 713

Менделя данные о горохе (Mendelian pea data), (пример 30.1) 563
Менденхолл (Mendenhall W.), цензурированные экспоненциальные выборки 705; библиография по испытаниям на долговечность 705
Меррингтон (Merrington M.), функции мощности критериев независимости в таблицах 2×2 743
Мёрфи (Murphy R. B.), графики толерантных интервалов 697
Метод моментов (method of moments), см. Моменты
Мешающий параметр (nuisance parameter) 279; устранение его (упражнение 30.18) 621
Мизес (R. von Mises), критерий согласия 603
Микки (Mickey M. R.), задача о двух средних 201
Микульский (Mikulski P. W.), эффективность ранговых критериев 672
Миллер Л. (Miller L. H.), таблицы критерия согласия Колмогорова 611
Миллер (Miller R. G., Jr.), оценивание с устраненным смещением 20
Минимальная граница дисперсии (МГД) (minimum variance bound(MVB)) 24—25; условие достижимости 25, 30; МГД в конкретных задачах (примеры 17.6—17.10) 26—27; МГД- и МД-оценки 27—28; улучшение 28—34; асимптотически достижимая МГД для любой функции от параметра 32; МГД и достаточность 42—43; МГД меньшая, чем граница дисперсии, когда неизвестно несколько параметров, (упражнение 17.20) 56; ослабление условий регулярности (упражнения 17.21—17.22) 56; аналоги (упражнение 17.23) 56; МГД- и МП-оценка 59—61, 68—70, (упражнение 18.4) 100; обобщение (упражнение 18.32) 108; асимптотические улучшения в нерегулярных случаях (упражнения 32.8—32.11) 713—714; при последовательном выборе 819—821
Минимальная достаточность (minimal sufficiency), см. Достаточность
Минимальной дисперсии (МД-) оценки (minimum variance (MV) estimators) 23—36; МД- и МГД-оценки 27; единственность 34—35; МД-оценки и достаточные статистики 44—45, (упражнение 17.24) 56—57; в классе несмещенных линейных комбинаций 115—116; МД- и МП-оценки 60—61, (упражнение 18.13) 103; единственность и полнота, 254—255; случай урезания и цензурирования 702—703
Минимума хи-квадрат (МХК-) оценки (minimum Chi-square estimators) 132—136; видоизмененные 133—134; асимптотическая эквивалентность МП-оценкам (упражнение 19.13) 138; среднее и дисперсия (упражнение 19.14) 138
Митра (Mitra S. K.), урезанное распределение Пуассона (упражнение 32.20) 716; модели для таблицы $r \times c$ 749; гипотезы в таблицах с несколькими входами 780
Многократные сравнения (multiple comparisons), в таблицах $r \times c$ 752
Многомерное нормальное распределение (multi-normal distribution), дисперсия квадратичной формы (упражнение 19.3) 136; случай, когда существует одномерная достаточная статистика без РНМ критерия (пример 22.11) 240—241; несмещенность критериев ОП о независимости 329; бисериальные методы (упражнение 26.12) 420; частная корреляция и регрессия 423—430, (упражнения 27.4, 27.6, 27.21) 457—458, 460; инвариантность свойства независимости при ортогональном преобразовании 444, (упражнение 27.7) 458; множественная корреляция 450—457; с матрицей внутриклассовой корреляции (упражнения 27.17—27.18) 459; предельные значения величины X^2 (упражнения 33.7—33.8) 782—783
Множественная корреляция (multiple correlation) 445—456; коэффициент 445—447; геометрическая интерпретация 447; отсеивание переменных 448—449; условное выборочное распределение в нормальном случае 449—450, (упражнение 27.13) 458; безусловное распределение в многомерном нормальном случае 450—454, (упражнения 27.14—27.16) 459; несмещенное оценивание в многомерном нормальном случае 456—457; м. к. с некоррелированными регрессорами (упражнение 27.20) 460
Мозес (Moses L. E.), критерии различия в масштабе 675
Молдон (Mauldon J. G.), фидуциальный парадокс 212
Моменты (moments), метод моментов, эффективность, 97—100, (упражнения 18.9, 18.17, 18.20, 18.27—18.28) 102—104, 106—107
Морэн (Moran P. A. P.), распределение множественного коэффициента корреляции 452
Мошмэн (Moshman J.), двойная выборка 826
Мощность (power), критерия 223—225; функция мощности 242; см. Критерии для проверки гипотез
МП, см. Максимального правдоподобия оценки
Муд (Mood A. M.), двухвыборочные критерии 674; критерий для проверки случайности, основанный на медиане (упражнение 31.12) 684; критические значения для критерия знаков 689
Мультиномиальное распределение (multinomial distribution), последовательная аппроксимация МП-оценки (пример 18.10) 77—79; МП-оценка числа классов, когда все вероятности одинаковы (упражнение 18.10) 102; эффективность критериев ОП 370; м. р. как основа критериев согласия 559—574; критерии согласия для данных о горохе (пример 30.1) 563; критерий однородности 788
Мур (Мооге А. Н.), цензурированные нормальные выборки 704; цензурированные логнормальные выборки 707
Мур П. (Mooге P. G.), оценивание, использующее две лучшие порядковые статистики (упражнение 32.14) 715
МХК, см. Минимума хи-квадрат оценки
Мэсси (Massey F. J., Jr.), таблицы критерия согласия Колмогорова 611; мощность критерия Колмогорова 613—616, (упражнение 30.16) 620

Наилучшие асимптотически нормальные (НАН) оценки (best asymptotically normal (BAN) estimators) 131—136
Наименьших квадратов (НК-) оценки (least squares (LS) estimators) 109—131, (упражнения 19.1—19.8) 136—137; эквивалентность с МП-оценками в нормальном случае 109—116; принцип НК 109—111;

линейная модель 111—126; несмещенность 113; матрица рассеяния 113; МД-свойство 115—116; геометрическая интерпретация 116—117; минимизация обобщенной дисперсии 117—118; оценивание дисперсии ошибки 118—120; роль предположения нормальности 120; случай вырождения 122—125; порядковое оценивание параметров сдвига и масштаба 126—131, (упражнения 19.10--19.12) 137—138; неэффективность в пуассоновском случае (упражнение 19.16) 138; общая линейная гипотеза 336; приближенная линейная регрессия 382—384; 434—435; НК-оценки в линейной регрессионной модели 472—493; использование дополнительной информации (упражнения 28.17—28.18) 496—497; уточнение для дополнительного наблюдения (упражнение 28.19) 497; двухэтапные (упражнение 28.24) 498—499; в $r \times c$ таблицах 760; см. Линейная модель, Регрессия
Найт (Knight W.), обратный выбор 795
НАН, см. Наилучшие асимптотически нормальные оценки
Нарайн (Narain R. D.), несмещенные критерии ОП для независимости 329
Независимость (independence), доказательства, использующие достаточность и полноту (упражнение 23.6—23.10) 296; н. и корреляция 373—374; критерии 633—647; н. и связь 721, сноска; частоты 722—723
Нейман (Neyman, J.), состоятельность МП-оценок 93, 514; НАН-оценки 131—133; доверительные интервалы 139, 154 (сноска); наиболее селективные интервалы 151—162; интервалы для правого конца равномерного распределения (упражнения 20.3—20.4) 179; критерии для проверки гипотез 219; максимальная мощность 224; проверка простой гипотезы H_0 против простой H_1 225; НКО критериев для нормальных параметров (пример 22.8) 237; РНМ критерии и достаточные статистики 239, (упражнение 22.11) 248; достаточность и подобные области 254; смещенность критерия для нормальной дисперсии (пример 23.12) 273; несмещенные критерии 277; достаточность, подобные области и независимость (упражнения 23.3—23.4) 295; метод ОП 301; таблицы функции мощности t-критерия 340; критерии ОП в k нормальных выборках (упражнения 24.4—24.6) 346—347; несущественные и структурные параметры 512; состоятельное оценивание структурной зависимости 514, 534; состоятельность критерия согласия X^2 579; «гладкий» критерий согласия 593—596; см. Типа А, A_1, В, B_1, С критерии
Неймана — Пирсона лемма (Neyman — Pearson lemma) 225, 229; обобщение 280—281
Непараметрические гипотезы (non-parametric hypotheses) 220—221; н. г. и свободные от распределения методы 629
Несмещенное оценивание (unbiassed estimation) 17—19; абсурдное (упражнение 17.26) 57; см. Смещение при оценивании
Несмещенные критерии (unbiassed tests) 271; н. к. и подобные критерии 275—276
Несущественные параметры (incidental parameters) 512
Нётер (Noether G. E.), доверительные интервалы для отношения биномиальных параметров (упражнение 20.9) 180; АОЭ 355; консервативность свободных от распределения процедур в дискретных случаях 667
Нецентральное F-распределение (non-central F-distribution) 337—341
Нецентральное t-распределение (non-central t-distribution) 341—342
Нецентральное χ^2-распределение (non-central χ^2-distribution) 305—310, 339, (упражнения 24.1—24.3) 345—346, (упражнение 24.19) 350; и АОЭ 368—369
Нецентральные доверительные интервалы (non-central confidence intervals) 144—145
НК, см. Наименьших квадратов оценки
НКО, наилучшая критическая область (BCR, best critical region) 225; см. Критерии для проверки гипотез
Нормальное распределение (normal distribution), оценивание среднего 15, 21; МГД для среднего (пример 17.6) 26; МГД для дисперсии (пример 17.10) 27; эффективность выборочной медианы (пример 17.12) 38; эффективность выборочного среднего отклонения (пример 17.13) 38—39; достаточность при оценивании среднего и дисперсии (пример 17.17) 48; границы для дисперсии при оценивании стандартного отклонения (упражнение 17.6) 54; несмещенная МД-оценка квадрата среднего (упражнение 17.7) 54; оценивание дисперсии с наименьшим средним квадратом ошибки (упражнение 17.16) 55; эффективность средней разности (упражнение 17.27) 57; несмещенная МП-оценка для среднего (пример 18.2) 62; МП-оценка стандартного отклонения (пример 18.8) 72; МП-оценка среднего и дисперсии (примеры 18.11, 18.13) 82, 86; оценивание среднего, принимающего только целочисленные значения (упражнение 18.21—18.23) 104—105; поправка на группировку в МП-оценках среднего и дисперсии (упражнение 18.25) 106; МП-оценка общего среднего для совокупностей с разными дисперсиями (упражнение 18.30) 107; МП-оценка среднего, функционально связанного с дисперсией (упражнение 18.31) 108; отсутствие МП-оценки (упражнение 18.34) 108; эквивалентность НК- и МП-оценок 109, 116; доверительные интервалы для среднего (примеры 20.1, 20.3, 20.5) 141—143, 148—149, 153—155; доверительные интервалы для дисперсии (пример 20.6) 156—158, (упражнение 20.6) 179, (упражнение 20.10) 180, (упражнение 23.18) 298; таблицы доверительных интервалов для дисперсии и отношения двух дисперсий 163; доверительные интервалы для среднего при неизвестной дисперсии 170—172; доверительные интервалы для отношения двух средних (пример 20.7) 172—174; доверительные области для среднего и дисперсии (упражнение 20.5) 179; доверительные интервалы для отношения двух дисперсий (упражнение 20.7) 180, (упражнение 23.18) 298; фидуциальные интервалы для среднего (пример 21.1) 187; фидуциальные интервалы для среднего при неизвестной дисперсии 188—189; доверительные и фидуциальные интервалы для задачи о двух средних 190—205, (упражнения 21.4—21.5, 21.9—21.10) 217, 218, (пример 23.10) 268—269, (пример 23.16) 288—290, (пример 24.2) 303—304; байесовские интервалы 205—207; проверка простой гипотезы H_0 о среднем (примеры 22.21—22.23) 223—224, 226—228, (примеры 22.12—22.14) 242—243, 247, (упражнение 22.12) 248, (пример 23.11) 270—

272, 285, (примеры 25.1—25.5) 353—368; проверка нормальной формы против двойной экспоненциальной (пример 22.5) 230—231; проверка различных гипотез о среднем и дисперсии (пример 22.8) 236—238; проверка простой гипотезы H_0 о дисперсии (упражнения 22.3, 22.5) 248, 285; отсутствие подобных областей (пример 23.1), 253; проверка сложной гипотезы H_0 о среднем (пример 23.7) 264—265, 277, (пример 23.14) 285—287, (пример 23.17) 293—294, (пример 24.1) 302—303; проверка сложной гипотезы H_0 о разности двух средних, когда дисперсии равны (пример 23.8) 265—266, (пример 23.15) 287—288; проверка сложной гипотезы H_0 для дисперсии (примеры 23.12—23.14) 272, 276, 285, (упражнение 23.13) 297, (примеры 24.3, 24.5) 311—313, 323—324; проверка сложной гипотезы H_0 о линейных функциях от различающихся средних и об общей дисперсии (пример 23.15) 287, (упражнение 23.19) 298; проверка сложной гипотезы H_0 для взвешенной суммы величин, обратных к различающимся дисперсиям (пример 23.16) 288; проверка сложной гипотезы H_0 о дисперсионном отношении (пример 23.16) 288, (упражнения 23.14, 23.17, 23.18) 297—298; доказательства свойств независимости, использующие полноту (упражнения 23.8, 23.9) 296; «особые» РНМН критерии (упражнения 23.16, 23.20) 297, 298; минимальность и одномерная достаточность (упражнение 23.31) 300; проверка равенства нескольких дисперсий (примеры 24.4, 24.6) 314, 328, (упражнения 24.4, 24.7) 346—347; в общей линейной гипотезе 333—344; таблицы функции мощности t-критерия 342; критерии ОП для k выборок (упражнения 24.4—24.6) 346—347; асимптотически наиболее мощные критерии для среднего (пример 25.1) 353; АОЭ выборочной медианы (примеры 25.2, 25.4, 25.5) 358, 364, 368; данные о функциональной зависимости (примеры 29.1—29.7) 518, 521, 529, 534, 540, 541, 543; непригодность метода смешанных семиинвариантов для оценивания функциональной зависимости 526—529, 548—549; проверка нормальности 558—559, 617; выбор классов для критерия X^2 574; «гладкий» критерий для среднего (пример 30.5) 598—599; устойчивость критериев 622—627; использование в критериях нормальных меток 651—653; использование двухвыборочного критерия 654—655, 657; эффективность критерия Вилкоксона и критерия нормальных меток 667, 671—672; парный t-критерий 680; урезание и цензурирование 703—705; критерии для выпадающих наблюдений 707—710; оценивание, использующее две наилучшие порядковые статистики (упражнение 32.14) 715; распределение экстремального отклонения от выборочного среднего (упражнение 32.21) 716—717; последовательный критерий для простой гипотезы H_0 о среднем (примеры 34.8, 34.10) 810, 813; последовательные критерии для дисперсии (пример 34.9) 812, (упражнения 34.18—34.20) 831; последовательный t-критерий 816—819; последовательное оценивание среднего (пример 34.12) 823; двойная выборка для среднего и дисперсии 824—826; распределение объема выборки при последовательном выборе (упражнения 34.12—34.13) 829—830; см. также Двумерное нормальное распределение, Многомерное нормальное распределение

Нормальные метки (normal scores) 651—653; в двухвыборочном критерии c_1 667—673, (упражнение 31.23) 686

Нортон (Norton H. W.), взаимодействие в таблицах с несколькими входами 781

Нулевая гипотеза (null hypothesis) 222, сноска

Нэйр (Nair K. R.), оценивание функциональной зависимости 538—539; таблицы доверительных интервалов для медианы 695; распределение экстремального отклонения от выборочного среднего (упражнение 32.21) 716—717; таблицы стьюдентизованного экстремального отклонения 710

Обобщенная дисперсия (generalized variance) 85; асимптотическая минимизация МП-оценками 85—86; минимизация в линейной модели НК-оценками 117—118

Обратный выбор (inverse sampling) 795

Обрубленные оценки (trimmed estimators) 704, 710

Общая линейная гипотеза (general linear hypothesis), см. Линейная модель

Огава (Ogawa J.), цензурированные экспоненциальные выборки 706

Огберн (Ogburn W. F.), данные о преступлениях в городах (пример 27.2) 442

Ограничения (constraints), налагаемые гипотезой 221, 249

Ограниченная полнота (bounded completeness), см. Полнота

Оделл (Odell P. L.), НК-оценка в случае вырождения 124

Однородности критерии (homogeneity tests) биномиальный и пуассоновский 774—776

Однородные категоризации (identical categorizations) 720

Однородные ошибки (identical errors) 468

Олкин (Olkin I.), несмещенное оценивание нормальной корреляции 393; несмещенное оценивание коэффициента множественной корреляции 456

ОП, см. Критерии отношения правдоподобия

Оперативная характеристика (operating characteristic) 798—799

Ортогональный (orthogonal), регрессионный анализ 474—475; о. полиномы в регрессии 475—480, (упражнение 28.23) 498; полиномы Лежандра 594 (сноска)

Остатки (residuals) 430 (сноска)

Отрицательно-биномиальное распределение (negative binomial distribution), МП-оценка (упражнения 18.26—18.27) 106; урезанное 707; последовательная выборка по качественному признаку (пример 34.1) 792

Оуэн (Owen D. B.), таблицы мощности t-критерия Стьюдента 342

ОХ(ОС), см. Оперативная характеристика

Оценивание (estimation), точечное 15—138; интервальное 139—218; о. и полнота 254—255; см. Эффективность оценивания

Оценка (правило) (estimator) и значение оценки (estimate) 14

Ошибки (errors), в модели НК 111; относительно регрессии 430; однородные 468; наблюдений и функциональных зависимостей 500—503; влияние ошибок наблюдений на регрессию 551—554

Пайк (Pyke R.), выборочные промежутки 687
Параметрические гипотезы (parametric hypotheses) 220
Параметры с ограничениями (restricted parameters), их МП-оценки (упражнения 18.21—18.23) 104—105
Парный t-критерий (paired t-test) 680
Патил (Patil G. P.), достижимость МГД 31
Патнайк (Patnaik P. B.), нецентральные χ^2- и F-распределения 308, 310, 339—341, (упражнение 24.2) 345—346; дисперсия величины X^2 (упражнение 30.5) 619; функция мощности критерия независимости в таблицах 2×2 743
Пейдж (Page E. S.), последовательный критерий в случае экспоненциального распределения (упражнение 34.14) 830
Перекрестное отношение (cross-ratio), в таблице 2×2 733, 744
Пересечение (intersection) полиномиальных регрессий 486
Перестановок критерии (permutation tests) 632
Пжиборовский (Przyborowski J.), доверительные границы для пуассоновского параметра 163
Пиллаи (Pillai K. C. S.), таблицы для стьюдентизованного экстремального отклонения 710
Пинкхэм (Pinkham R. S.), моменты нецентральной t-величины 342
Пирс (Peers H. W.), байесовские и доверительные интервалы 215
Пирсон (Pearson E. S.), графики доверительных интервалов для биномиального распределения 162; критерии для проверки гипотез 219; максимальная мощность 224; проверка простой гипотезы H_0 против простой альтернативы H_1 225; НКО критериев для нормальных параметров (пример 22.8) 237; РНМ критерии и достаточные статистики 239; выбор размера критерия 247; РНМ критерий и достаточность в случае экспоненциального распределения (упражнение 22.11) 248; смещенность критерия для нормальной дисперсии (пример 23.12) 273; несмещенные критерии 277; метод ОП 301; графики функции мощности для линейной гипотезы 340; критерии ОП при k нормальных выборках (упражнения 24.4—24.6) 346—347; критерии согласия 595, 604, (упражнение 30.12) 620; исследования устойчивости 623; выпадающие наблюдения 708; таблицы стьюдентизованного интервала 709; модели для таблицы 2×2 736; функции мощности критериев в таблицах 2×2 743
Пирсон К. (Pearson K.), развитие корреляционной теории 374—375; тетрахорическая корреляция 410; бисериальная корреляция (таблица 26.7) 411, 413, 465; критерий согласия X^2 559; оценивание, использующее две наилучшие порядковые статистики (упражнение 32.14) 715; коэффициент сопряженности 746; стандартная ошибка величины X^2 в таблицах $r \times c$ 750—751; таблицы с несколькими входами 776, (упражнение 33.7) 782
Пирсона распределения (Pearson distributions), эффективное оценивание для них 97—98, (пример 18.18) 98, (упражнение 18.16) 103
Питмэн (Pitman E. J. G.), «близкие» оценки 22; распределения, обладающие достаточными статистиками 46, 48, 51; доверительные интервалы для отношения дисперсий в двумерном нормальном распределении (упражнение 20.19) 182; достаточность статистики ОП для проверки простой гипотезы H_0 (упражнение 22.13) 248; несмещенные инвариантные критерии для параметров сдвига и масштаба 325, (упражнение 24.15) 349; АОЭ 355; АОЭ критерия знаков (упражнение 25.2) 371, 692; АОЭ критерия Вилкоксона (упражнения 25.3—25.4) 371, 665, (упражнение 31.22) 686; критерий независимости 635; двухвыборочный критерий 655—657; состоятельность критерия Вилкоксона 663—664, (упражнение 31.16) 685
ПКОВ (SPR tests), последовательные критерии отношения вероятностей 801—814, (упражнения 34.4, 34.7—34.11, 34.21—34.22) 828—829, 831—832
Плэкетт (Plackett R. L.), об истоках теории НК 115; теория НК в случае вырождения 122, (упражнения 19.6—19.8) 136—137; общая линейная модель (упражнение 19.2) 136; поправки на непрерывность 682; 745; упрощенное оценивание в цензурированных выборках 703; цензурированные нормальные выборки 704; урезанное пуассоновское распределение 706; (упражнение 32.20) 716; х. ф. и семиинварианты логистических порядковых статистик (упражнение 32.12) 714; оценивание квантилей (упражнение 32.13) 714—715; взаимодействие в таблицах с несколькими входами 781
Погода и урожаи (weather and crops), данные Хукера (пример 27.1) 439
Подобные области (similar regions), подобные критерии 251—252; см. Критерии для проверки гипотез
Полиномиальная регрессия (polynomial regression) 475—482
Политомия (polytomy) 730
Полихорическое оценивание нормальной корреляции (polychoric estimation of normal correlation) 752
Полнота (completeness) 254; п. и единственность несмещенной оценки 255; п. достаточных статистик 254—259; п. и подобные области 263; в непараметрических задачах 631—632; последовательная 819 (сноска)
Полсон (Paulson A. S.), критерии для выпадающих наблюдений 710
Полсон (Paulson E.), критерии ОП для экспоненциальных распределений 329, (упражнения 24.16, 24.18) 349—350; выпадающие наблюдения 709
Поправки (corrections), на смещение при оценивании 19—22, (упражнения 17.17—17.19) 55; на группировку в МП-оценках (упражнения 18.24—18.25) 105—106
Поправки на непрерывность (continuity corrections) в свободных от распределения критериях 681—683; в таблицах 2×2 744
Порядковые статистики (order-statistics), для параметра сдвига Коши (пример 18.9) 76; в НК-оценках параметров сдвига и масштаба 126—131; п. с. и подобные области 252; полнота 631—632; использование нормальных меток 651—653; асимптотическое распределение математических ожиданий 653; критерий знаков для квантилей 687—692; доверительные интервалы для квантилей 692—695; толерантные интервалы 695—698; при точечном оценивании 699; урезание и цензуриро-

вание 699—707; выпадающие наблюдения 707—712; оценивание квантилей (упражнение 32.13) 714—715
Последовательные методы (sequential methods) 791—832; для качественных признаков (упражнение 18.18) 104, 791—810; замкнутые, открытые и усеченные схемы 792; последовательные критерии 798; ОХ 798—799; СОВ 799—800; ПКОВ 801—814, (упражнения 34.4, 34.7—34.11) 828—829; правила остановки при обследовании с заменой всех дефектных изделий годными и при двойной выборке 809—810; эффективность 812—814; сложные гипотезы 814—816; последовательный t-критерий 816—819; оценивание 819—824, (упражнения 34.5—34.6) 828; двойная выборка 824—826, (упражнения 34.15—34.17) 830—831; свободные от распределения 826
Поттхофф (Potthoff R. F.), проверка однородности 776
Правдоподобия принцип (likelihood principle) 58
Правдоподобия уравнение (likelihood equation) 59; см. Максимального правдоподобия оценки
Правдоподобия функция (ФП) (likelihood function (LF)) 23; использование 93, см. Максимального правдоподобия оценки
Прайс (Price R.), нецентральное F-распределение 339
Праут (Prout T.), см. Левонтин
Предсказания интервалы (prediction intervals) 484
Преобразования (transformations), функциональных соотношений в линейные 550; нормализующие 627—628
Пресс (Press S. J.), сравнения с помощью доверительных интервалов в задаче о двух средних 202
Преступность в городах (crime in cities), данные Огберна (пример 27.2) 442—443
Простая гипотеза (simple hypothesis) 218—248; см. Гипотезы, Критерии для проверки гипотез
Прэтт (Pratt J. W.), ожидаемая длина доверительных интервалов (упражнения 20.17—20.18) 181; несмещенное оценивание нормальной корреляции 393; несмещенное оценивание множественной корреляции в многомерном нормальном случае 456; устойчивость двухвыборочных критериев 673
Прямоугольное распределение (uniform distribution), см. Равномерное распределение
Псевдообратная матрица (pseudo-inverse matrix) (упражнение 27.21) 460
Пуассоновское распределение (Poisson distribution), МГД для параметра (пример 17.8) 26; абсурдная несмещенная оценка в усеченном случае (упражнение 17.26) 57; МХК-оценка (пример 19.11) 134; п. р. в линейной модели (упражнение 19.16) 138; доверительные интервалы (пример 20.4) 149—151; таблицы доверительных интервалов 163; проверка простой гипотезы H_0 (упражнение 22.1) 247, 285; РНМН критерии для разности между двумя пуассоновскими параметрами (упражнение 23.21) 298; п. р. и нецентральное χ^2-распределение (упражнение 24.19) 350; урезание и цензурирование 706, (упражнения 32.19—32.20, 32.22—32.24) 716—717; дисперсионный критерий 775—776; обратный выбор 795; последовательное оценивание (пример 34.13) 824
Пури (Puri M. L.), k-выборочные критерии 676
Путтер (Putter J.), исследование совпадений 682
Пэбст (Pabst M. R.), эффективность критерия ранговой корреляции 646
Пэчерс (Pachares J.), таблицы доверительных границ для параметра биномиального и дисперсии нормального распределений 163; таблицы стьюдентизованного интервала 710

Равномерно наиболее мощный (РНМ) (uniformly most powerful (UMP)) 234; см. Критерии для проверки гипотез
Равномерно наиболее мощный несмещенный (uniformly most powerful unbiassed (UMPU)) 272; см. Критерии для проверки гипотез
Равномерное распределение (uniform distribution, rectangular distribution), наибольшее наблюдение как достаточная статистика для верхней границы (пример 17.16) 43; достаточность, когда обе границы зависят от параметров (примеры 17.18, 17.21—17.23) 51—52; несмещенная МД-оценка (упражнение 17.24) 56—57; смещенность МП-оценки для границы интервала (пример 18.1) 61; МП-оценка, являющаяся функцией только от одной из двух достаточных статистик (пример 18.5) 67; МП-оценки обеих границ (пример 18.12) 83; неединственность МП-оценки для параметра сдвига (упражнение 18.17) 103—104; порядковые НК-оценки параметров масштаба и сдвига (пример 19.10) 129; доверительные интервалы для верхней границы (упражнения 20.2—20.4, 20.15) 179, 181; РНМ односторонние критерии для простой гипотезы H_0 о параметре сдвига (упражнение 22.8) 248; минимальная достаточность (упражнение 23.12) 297; мощность условного критерия (упражнение 23.23) 299; РНМ критерий для сложной гипотезы H_0 о параметре сдвига (упражнение 23.26) 299; связь с χ^2-распределением 317—318; критерий ОП для параметра сдвига (упражнение 24.9) 347; критерий ОП для параметра масштаба (упражнение 24.16) 349; распределение порядковых статистик (упражнение 30.17) 620—621; АОЭ двухвыборочного критерия Вилкоксона (упражнение 31.22) 686; асимптотическая граница дисперсии оценки среднего (упражнение 32.10) 714; оценивание для цензурированных выборок (упражнение 32.25) 718
Рагхавачари (Raghavachari M.), критерии сдвига 674
Радж (Raj D.), урезанные и цензурированные гамма-распределения 706
Разбиения величины X^2 (partitions of X^2) 601—602, в таблицах $r \times c$ 769—774
Размер критерия (size of test), 222; его выбор 245—247
Разорение игрока (gambler's ruin), (пример 34.2) 796
Райдер (Rider P. R.), урезанные распределения 707
Рамачандрамурти (Ramachandramurty P. V.), устойчивое оценивание 673
Рамачандран (Ramachandran K. V.), таблицы доверительных границ для нормальной дисперсии и для отношения дисперсий 163, 274, (упражнение 23.17) 298

Ранги (ranks), как инструментальные переменные 540—543, (упражнения 29.9—29.10) 555; см. Ранговые критерии
Ранговые критерии (rank tests) 637—638; использование коэффициентов ранговой корреляции 637—651; для задачи двух и k выборок 659, 678; независимость симметрических функций (упражнение 32.2) 712—713
Рао Б. (Rao B. R.), аналог МГД (упражнение 17.23) 56; цензурирование, урезание и эффективность 702, (упражнение 32.26) 718
Рао (Rao C. R.), МГД 24, (упражнение 17.20) 56, достаточность и МД-оценка 44, 48; эффективность и корреляция 70; МХК-оценка 133; эффективность второго порядка 136
Расположение (location), см. Сдвига и масштаба параметры
Раштон (Rushton S.), ортогональные полиномы 480
Регрессия (regression) и зависимость 373—374; кривые р. 377—378; р. и ковариация 378—380; линейная р. 380—382; коэффициенты р. 381; уравнения р. 381; вычисление (примеры 26.6—26.7) 386—389; диаграмма разброса 389; стандартные ошибки 390—392; критерии регрессии и критерии независимости 395—396; линейность, коэффициент корреляции и корреляционное отношение 396—400; критерий ОП для линейности р. 400—403; общее 461—499; аналитическая теория 461—472, (упражнения 28.1—28.5) 493—494; критерии линейности 467—469; общая линейная модель 472—493; поправки 474, (упражнение 28.20) 497; кусочно-линейные кривые 474; ортогональная р. 474—482, (пример 28.4) 485, (упражнение 28.23) 495; доверительные интервалы и критерии для параметров 483—487, (упражнения 28.15—28.22) 495—498; доверительная область для линии р. 487—493; критерии для разности между коэффициентами ρ_i (упражнения 28.14—28.15) 495; использование дополнительной информации (упражнения 28.17—28.18) 496—497; графическая подгонка (упражнение 28.19) 497; р. и функциональная зависимость 502—503, 504—506; р. и контролируемые переменные 544—545; влияние ошибок наблюдений 551—554; см. также Корреляция, Линейная модель, Множественная корреляция, Наименьших квадратов оценки, Частная корреляция и регрессия
Регрессор (regressor) 461
Резников (Resnikoff G. J.), таблицы нецентрального t-распределения 342
Решающие функции (decision functions) 827
Рикер (Ricker W. E.), таблицы доверительных интервалов для пуассоновского параметра 163
РНМ, см. Равномерно наиболее мощный
РНМН, см. Равномерно наиболее мощный несмещенный
Роббинс (Robbins H.), границы для дисперсий оценок 34; оценивание стандартного отклонения нормального распределения (упражнение 17.6) 54; t-распределение Стьюдента, когда средние наблюдений различны 625; толерантные интервалы для порядковых статистик 695; моменты в последовательном оценивании 820
Робинсон (Robinson D. E.), пересечение полиномиальных регрессий 486
Робсон (Robson D. S.), оценивание границы области изменения (упражнение 17.13) 55; ортонормальные полиномы (упражнение 28.23) 498
Розенблатт (Rosenblatt M.), распределение критерия Гёдинга 677
Розенталь (Rosenthal I.), выборочные эксперименты с G в упорядоченных таблицах 758
Рой (Roy A. R.), выбор классов для критерия X^2 575
Рой Дж. (Roy J.), урезанное пуассоновское распределение (упражнение 32.20) 716
Рой К. (Roy K. P.), асимптотическое распределение статистики ОП 310
Рой С. (Roy S. N.), совместные доверительные интервалы 175; модели для таблиц $r \times c$ 749; гипотезы для таблиц с несколькими входами 780
Ромиг (Romig H. G.), двойная выборка 810
Ротенберг (Rothenberg T. J.), оценивание параметра сдвига распределения Коши 77

Салех (Saleh A. K. M. E.), цензурированные экспоненциальные выборки 706
Сандрум (Sundrum R. M.), оценивание эффективности и мощности 231, (упражнение 22.7) 248; мощность критерия Вилкоксона 667
Санкаран (Sankaran M.), граница дисперсии (упражнение 17.23) 56
Сархан (Sarhan A. E.), порядковые НК-оценки параметров сдвига и масштаба (упражнения 19.11—19.12) 137—138, (упражнение 32.18) 716; порядковые статистики 687; цензурированные выборки 704—706, (упражнение 32.25) 718
Свами (Swamy P. S.), эффективность при урезании и цензурировании 702—703
Свед (Swed F. S.), таблицы критерия серий (упражнение 30.8) 619
Сверхэффективность (supereffeciency) 70
Свободные от параметра критерии согласия (parameter-free tests of fit) 591—593, 605, 617—618; (упражнение 30.10) 620
Свободные от распределения процедуры (distribution-free procedures), критерии согласия 591—593, 603—604; доверительные интервалы для непрерывной ф. р. 611; общие сведения 628—629; непараметрические задачи 629; классификация 629—630; построение критериев 631—632; эффективность критериев 632—633; критерии независимости 633—647; критерии случайности 647—653; двухвыборочные критерии 653—675; доверительные интервалы для параметра сдвига 658, (упражнение 31.17) 685; k-выборочные критерии 675—679; критерии симметрии 679—681; критерии для квантилей 687—692; доверительные интервалы для квантилей 692—695; толерантные интервалы 695—698; критерии для выпадающих наблюдений 711—712; категоризованные данные 719—790; последовательные критерии 826—827
Связь (association), в таблицах 2×2 719—725; с. и независимость 721 (сноска); в таблицах $r \times c$ 752—769; частная 725—730, 777—781; см. Категоризованные данные
Сдвига и масштаба параметры (location and scale parameters), МП-оценки 93—95; некоррелированность МП-оценок в случае симметричного распределения 95; порядковые НК-оценки 126—131, (упражнения

19.10—19.12) 137—138; некоррелированность порядковых НК-оценок в случае симметричного распределения 129; полнота 255; отсутствие критерия для сдвига с мощностью, не зависящей от масштабного параметра (упражнение 23.30) 300; несмещенные инвариантные критерии для них 324—329, (упражнение 24.15) 349; оценивание при проверке согласия 593—605, (упражнение 30.10) 620
Серий критерий (runs tests) (упражнения 30.7—30.9) 619—620
Сечение (array) 377
Сиддики (Siddiqui M. M.), цензурированные экспоненциальные выборки 706
Сиджел (Siegel S.), двухвыборочный критерий для различия в масштабе 674—675
Сил (Seal H. L.), критерий серий как дополнение к критерию согласия X^2 591, (упражнение 30.7) 619
Силбайндер (Seelbinder B. M.), двойная выборка 826
Силверстоун (Silverstone H.), МГД 24; сравнение НАН-оценок 136
Силлито (Sillito G. P.), порядковые статистики 708
Симметрии критерии (symmetry tests) (упражнение 31.25) 679—681
Симметричные распределения (symmetrical distributions), некоррелированность МП-оценок параметров сдвига и масштаба 95—96; некоррелированность порядковых НК-оценок параметров сдвига и масштаба 129; условие на НК-оценку параметра сдвига для того, чтобы она была выборочным средним (упражнение 19.10) 137; АОЭ критерия знаков 690—692, (упражнение 32.1) 712; критерии симметрии 679—681
Сингх (Singh R.), полнота 255
Сичел (Sichel H. S.), МП-оценки в логнормальном распределении (упражнения 18.7—18.9) 101—102
Скедастическая кривая (scedastic curve) 462
Скотт (Scott E. L.), состоятельность МП-оценок 93, 514, 534; несущественные и структурные параметры 512
Сложные гипотезы (composite hypotheses) 249—300; см. Гипотезы, Критерии для проверки гипотез
Сложные таблицы сопряженности (complex contingency tables), см. Таблицы с несколькими входами
Случайность (randomness) критерии случайности 647—651
Слэктер (Slakter M. J.), аппроксимация X^2 588
Смеси распределений (mixtures of distributions), МГД для пропорций (упражнение 17.2) 53
Смещение при оценивании (bias in estimation) 17—19; поправки на смещение 19—22, (упражнения 17.13, 17.17—17.18) 55; см. Несмещенное оценивание
Смещение критериев (bias in tests) 270; см. Несмещенные критерии, Критерии для проверки гипотез
Смирнов Н. В., критерий согласия 603; таблицы критерия согласия Колмогорова 610; односторонний критерий согласия 612—613; двухвыборочный критерий 673
Смит (Smith C. A. B.), моменты величины X^2 в таблице $r \times c$ (упражнение 33.9) 783
Смит Дж. (Smith J. H.), МП-оценки групповых вероятностей в таблице $r \times c$ (упражнение 33.29) 789
Смит К. (Smith K.), МХК-оценки для пуассоновского распределения (пример 19.11) 135
Смит С. (Smith S. M.), семиинварианты МП-оценки 72
Смит У. (Smith W. L.), урезание и достаточность 700
Со (Saw J. G.). цензурированные нормальные выборки 705
Собел (Sobel M.), свойство независимости экспоненциального распределения (упражнение 23.10) 296; цензурирование в экспоненциальных выборках 705, (упражнение 32.17) 716
СОВ, см. Средний объем выборки
Совпадения (ties) и свободные от распределения критерии 682—683
Сопер (Soper H. E.), стандартная ошибка коэффициента бисериальной корреляции 415—416
Сопряженность (contingency), таблицы 745; коэффициент 746; см. Категоризованные данные
Состоятельность критериев (consistency in tests) 322; критериев ОП 323
Состоятельность при оценивании (consistency in estimation) 15—17, 132 (сноска), 351; МП-оценок 64—67, 83, (пример 18.16) 92, (упражнение 18.35) 108
Спёрджен (Spurgeon R. A.), таблицы нецентрального t-распределения 342
Спрент (Sprent P.), линейные функциональные зависимости 526
Спротт (Sprott D. A.), фидуциальный и байесовский методы 215
Срединная корреляция (medial correlation) (упражнение 31.7) 683
Средний квадрат ошибки (mean-square-error), оценивание по методу минимума с. к. о. 39—40, (упражнение 17.16) 55
Средний объем выборки (average sample number) 799—800
Статистическая зависимость (statistical relationship) 373
Стейн (Stein C.), оценивание параметров с ограничениями (упражнение 18.22) 105; неподобные критерии для сложной гипотезы H_0 269; бесполезность критерия ОП (пример 24.7) 330; инвариантность и достаточность 344; критерии симметрии 679; ПКОВ 801; последовательная полнота 819 (сноска); двойная выборка 824—826
Степени свободы (degrees of freedom), гипотезы 221, 249
Стерн (Sterne T. E.), доверительные интервалы для пропорции 163
Стефан (Stephan F. F.), МП-оценки групповых вероятностей в таблице $r \times c$ (упражнение 33.29) 789
Стефенс (Stephens M. A.), критерии согласия 604
Стивенс (Stevens W. L.), доверительные интервалы в дискретном случае 164; фидуциальные интервалы для биномиального параметра (упражнение 21.3) 216—217; критерий серий (упражнение 30.8) 619
Структурная зависимость (structural relationship), см. Функциональные и структурные зависимости
Структурные параметры (structural parameters) 512
Стьюарт (Stuart A.), оценивание коэффициента корреляции в нормальном случае (упражнение 18.12) 103; АОЭ и максимальная потеря мощности 366; многомерное нормальное распределение с внутри-

классовой корреляцией (упражнения 27.18—27.19) 459—460; критическая область для критерия согласия X^2 562; совместное распределение ранговых корреляций 644; АОЭ критериев случайности 653, (упражнения 31.10—31.12) 684; корреляция между рангами и значениями варианты (упражнения 31.13—31.14) 684—685; проверка гипотезы о разности двух коэффициентов связи 757; критерий однородности маргинальных распределений в таблицах $r \times r$ (упражнение 33.24) 788

Стьюдент («Student», W. S. Gosset), стьюдентизация 171; выпадающие наблюдения 708

Стьюдентизация (studentization) 170—172, (пример 20.7) 172—174

Стьюдентизованное максимальное абсолютное отклонение (studentized maximum absolute deviate) 709

Стьюдентизованное экстремальное отклонение (studentized extreme deviate) 709

Стьюдентизованный размах (studentized range) 710

Субрахманиам (Subrahmaniam K.), урезанное пуассоновское распределение (упражнение 32.20) 716

Сукхатм (Sukhatme P. V.), критерии ОП для k экспоненциальных распределений (упражнения 24.11—24.13) 348—349

Сходимость по вероятности (convergence in probability) 16; см. Состоятельность при оценивании

Сэвидж (Savage I. R.), библиография о свободных от распределения статистиках 630; двухвыборочный критерий с нормальными метками 668, 672; независимость симметричных и ранговых статистик (упражнение 32.2) 712—713

Сэвидж Л. (Savage L. J.), достаточность 42

Сэмпфорд (Sampford M. R.), урезанное отрицательно-биномиальное распределение 707

Сюй (Hsu P.), критерий независимости в таблицах 2×2 741, 743

Сюй (Hsu P. L.), оптимальное свойство критерия ОП для линейной гипотезы 344

Таблицы 2×2 (2×2 tables), связь в них 720—725, (упражнения 33.1—33.2) 781; частная связь 725—730; вероятностная интерпретация мер связи 730—733; асимптотические критерии независимости 733—736; точные критерии независимости для разных моделей 736—744; поправка на непрерывность 744—745, (упражнение 33.5) 782; компоненты величины X^2 в таблицах 2×2×2 777, (упражнения 33.27—33.28) 788—789; критерий однородности маргинальных распределений в таблицах 2^c (упражнение 33.30) 790

Таблицы с несколькими входами (multiway tables) 776—781

Тагути (Taguti G.), таблицы толерантных границ для нормального распределения 178

Тайкью (Tiku M. L.), нецентральные χ^2- и F-распределения 307, 339, 341; критерии согласия 605

Тамура (Tamura R.), критерии масштаба 674

Тартер (Tarter M. E.), логистические порядковые статистики (упражнение 32.12) 714

Тейл (Theil H.), оценивание функциональной зависимости 539, 541, 549, (упражнения 29.9—29.10) 555

Тейнис (Tanis E. A.), проверка гипотез для экспоненциальных распределений (упражнение 24.13) 349

Тейт (Tate R. F.), несмещенная МД-оценка (упражнение 17.24) 57, таблицы доверительных интервалов для нормальной дисперсии 163; кратчайшие интервалы (упражнение 20.10) 180; бисериальная и точечно-бисериальная корреляция 416—417, (упражнения 26.10—26.12) 419—420; урезанное пуассоновское распределение 706, (упражнение 32.19) 716

Тейчер (Teicher H.), характеризация посредством формы МП-оценки (упражнение 18.2) 100; моменты в последовательном оценивании 820

Терпстра (Terpstra T. J.), k-выборочный критерий 678

Терри (Terry M. E.), критерии, использующие нормальные метки 652, 668

Тетрахорическая корреляция (tetrachoric correlation) 407—411; ряды 409, (упражнения 26.5—26.6) 418—419; ряды и канонические корреляции 763—764

Тиензо (Tienzo B. P.), стьюдентизованное экстремальное отклонение 710

Тиланус (Tilanus C. B.), оценивание параметра сдвига распределения Коши 77

Тинджи (Tingey F. H.), односторонний критерий согласия 613

Типа A, A_1, B, B_1, C критерии (type A, A_1, B, B_1, C tests) 277

Типа A распределение, зависящее от случайных параметров (type A contagious distribution), МП-оценки и оценки по методу моментов (упражнение 18.28) 107

Типа M несмещенные критерии (Type M unbiassed tests) 272

Типа I, Типа II цензурирование (type I, type II censoring) 700—701

Типа IV (Пирсона) распределение (type IV (Pearson) distribution), центр расположения и эффективность метода моментов (упражнение 18.16) 103

Тодхантер (Todhunter I.), применение Арбетнотом критерия знаков 687 (сноска)

Токарска (Tokarska B.), таблицы функции мощности t-критерия Стьюдента 342

Толерантные интервалы (tolerance intervals), для нормального распределения 176—179, (упражнения 20.20——20.22) 182; свободные от распределения 695—698, (упражнения 32.6—32.7) 713

Толерантные области (tolerance regions) 698

Томпсон (Thompson W. R.), доверительные интервалы для медианы 695; выпадающие наблюдения 709

Точер (Tocher K. D.), точный оптимальный критерий для таблиц 2×2 742

Точечно-бисериальная корреляция (point-biserial correlation) 416—417, (упражнение 26.5) 418

Точечное оценивание (point estimation) 13—138

Тренд (trend) 647

Трикетт (Trickett W. H.), таблицы для задачи о двух средних 202

Тьюки (Tukey J. W.), фидуциальный парадокс 212; двухвыборочный критерий против альтернативы об изменении масштаба 675; доверительные интервалы для квантилей 694; толерантные области 698; урезание и достаточность 700; выпадающие наблюдения 710

Тэнг (Tang P. C.), нецентральное F-распределение и таблицы функции мощности для линейной гипотезы 339—341; нецентральное χ^2-распределение (упражнение 24.1) 345
Тэтчер (Thatcher A. R.), биномиальные прогнозы 215

Уайз (Wise M. E.), аппроксимации для X^2 561
Уайт (White C.), таблицы критерия Вилкоксона 662
Уайтекер (Whitaker L.), данные о смертности женщин 135
Уиддер (Widder D. V.), преобразование Лапласа 256 (сноска)
Уикселл (Wicksell S. D.), аналитическая теория регрессии (упражнение 26.2) 418, 465, (упражнение 28.4) 494; критерий линейности регрессии 468
Уилкс (Wilks S. S.), кратчайшие доверительные интервалы 159; наименьшие доверительные области 175; доверительные интервалы для верхней границы равномерного распределения (упражнение 20.2) 179; асимптотическое распределение статистики ОП 310; обзор литературы о порядковых статистиках 689; толерантные интервалы 695, 698
Уилльямс (Williams C. A., Jr.), критерий согласия χ^2 587, 615
Уилльямс Э. (Williams E. J.), сравнение линейных моделей (упражнение 28.22) 498; основанные на метках методы для таблиц $r \times c$ 760; канонический анализ (упражнение 33.13) 785
Уиткомб (Whitcomb M. G.), несмещенные критерии для нормальной дисперсии 286
Уитлок (Whitlock J. H.), оценивание границы области изменения (упражнение 17.13)55
Уитни (Whitney D. R.), критерий Вилкоксона 660
Уиттинг (Witting H.), АОЭ критерия Вилкоксона 667; АОЭ критерия знаков 692
Уиттингхилл (Whittinghill M.), проверка однородности 776
Уишард (Wishart J.), нецентральные χ^2 и F-распределения 307, 339, (упражнение 24.1) 345; моменты множественного коэффициента корреляции 454; (упражнения 27.10—27.11) 458
Уокер (Walker A. M.), эффективность оценок 70
Уоллес Д. (Wallace D. L.), асимптотические разложения в задаче о двух средних 202
Уоллес (Wallace T. D.), двухэтапный метод НК (упражнение 28.24) 499
Уоллис (Wallis W. A.), толерантные границы 178; критерий Вилкоксона 660—662, (упражнение 31.15) 685; k-выборочный критерий 676; исследование совпадений 682
Уолш (Walsh J. E.), мера эффективности критерия 369; влияние внутриклассовой корреляции на t-критерий Стьюдента (упражнение 27.17) 459; мощность критерия знаков 692; толерантные интервалы 698; цензурированные нормальные выборки 705; выпадающие наблюдения 712
Уоркинг (Working H.), доверительная область для линии регрессии 487, 491—492, (упражнение 28.13) 495
Уормлитон (Wormleighton R.), толерантные области 698
Упорядоченная альтернатива (ordered alternative), для k-выборочного критерия 678
Упорядоченная категоризация (ordered categorization) 719
Упорядоченные таблицы $r \times c$ (ordered $r \times c$ tables) 752—769
Урезанные распределения (truncated distributions) 699—707; оценивание точки урезания (упражнение 17.13) 55
Урожаи пшеницы и картофеля (yields of wheat and potatoes), данные о них (таблица 26.3) 388
Усеченные последовательные схемы (truncated sequential schemes) 792
Условности принцип (conditionality principle) 292—293
Условные критерии (conditional tests) 293—295
Устойчивость (robustness) 622—627; при оценивании 627, 673
Уэлч (Welch B. L.), задача о двух средних 199—202; байесовские и доверительные интервалы 215; мощность условных критериев 293—295, (упражнение 23.23) 299; центральное t-распределение 342

Феллер (Feller W.), подобные области 252—254, (упражнения 23.1—23.2) 295; распределение статистики Колмогорова 605
Фергюсон (Ferguson T. S.), выпадающие наблюдения 708—709
Фередей (Fereday F.), линейные функциональные зависимости с несколькими переменными 526, (упражнения 29.6—29.8) 554—555
Ферон (Féron R.), характеризация двумерной нормальности 471 (сноска)
Фехнер (Fechner G. T.), работа по ранговой корреляции 640
Фидуциальные интервалы (fiducial intervals), фидуциальная вероятность, см. Фидуциальные выводы
Фидуциальные выводы (fiducial inference), общее 183—188; для распределения Стьюдента 188—189, (упражнение 21.7) 217; в задаче о двух средних 203—205; критическое обсуждение 207—216; парадоксы 211—212; связь с байесовским выводом 212—215
Фикс (Fix E.), таблицы нецентрального χ^2-распределения 310; линейность регрессии 553; таблицы критерия Вилкоксона 662
Филлер (Fieller E. C.), трудности теории доверительных интервалов 152, 173
Финни (Finney D. J.), МП-оценка в логнормальном распределении (упражнения 18.7—18.9, 18.19—18.20) 101—102, 104; урезанное биномиальное распределение 707; таблицы для точного критерия независимости в таблице 2×2 740—741; обратный выбор 795
Фишер (Fisher F. M.), оценивание параметра сдвига в распределении Коши 77
Фишер (Sir Ronald A. Fisher), определение ФП 23 (сноска); достаточность 40; разбиение ошибки при неэффективной оценке (упражнение 17.11) 55; принцип МП 58; последовательное приближение к МП-оценке 75—76, (пример 18.10) 77; использование ФП 93; МП-оценки параметров сдвига и масштаба 93; эффективность метода моментов 98, (упражнение 18.16)

103, (упражнение 18.27) 106; центр расположения для распределения типа IV (упражнение 18.16) 103; состоятельность 132 (сноска), 351; фидуциальные выводы 183; фидуциальное решение задачи о двух средних 204; фидуциальные парадоксы 211; фидуциальные интервалы для последующих наблюдений (упражнение 21.7) 217; исчерпывающая и достаточная статистика 260, сноска; вспомогательные статистики 292—293; нецентральные χ^2 и F-распределения 307, 339; преобразование коэффициента корреляции 394; критерии для корреляции и регрессии 400; распределение внутриклассового коэффициента корреляции 407, (упражнение 26.14) 420—421; распределение частных коэффициентов корреляции 444; распределение множественного коэффициента корреляции 452—453, (упражнения 27.13—27.16) 458—459; ортогональные полиномы 479—480; критерий различия между регрессиями (упражнение 28.15) 495; распределение критерия согласия X^2 565, 570 (упражнение 30.1) 618; критерии, использующие нормальные метки 652, 673; критерий симметрии 681; точное исследование таблицы 2×2 737, 739; дисперсионные критерии 775

Фокс (Fox K. A.), доверительные интервалы для многомерной нормальной множественной корреляции 457

ФоксЛ.(Fox L.), обращение матрицы 114

Фокс М. (Fox M.), графики функции мощности для линейной гипотезы 340

ФП, см. Правдоподобия функция

Френд (Freund R. J.), двухэтапный метод НК (упражнение 28.24) 498—499

Фреше (Fréchet M.), линейные регрессии и корреляционные отношения (упражнение 26.24) 422

Фримэн (Freeman G. H.), точный критерий независимости в таблицах $r \times c$ 749

Функциональные и структурные зависимости (functional and structural relations) 373; обозначения 501; линейные 501—503; ф. и с. з. и регрессия 501—505; МП-оценки 506—518, 545—547; геометрическая интерпретация 514—515, 547; доверительные интервалы и критерии 518—523; несколько переменных 523—526; использование смешанных семиинвариантов 526-530, 547—549; инструментальные переменные 530—545; использование рангов 540—543; контролируемые переменные 544—545, 549; криволинейные зависимости 545—550

Функция распределения (distribution function), выборочная 602—603; доверительные границы для ф. р. 611—613

Фуржо (Fourgeaud C.), характеризация двумерной нормальности 471 (сноска)

Фрэзер (Fraser D. A. S.), толерантные интервалы для нормального распределения 179; фидуциальные выводы 212, 215; минимальная достаточность 260 (сноска); критерий серий в дополнение к критерию согласия X^2 591, (упражнение 30.7) 619; ранговые критерии 653; толерантные области 698

Хадсон (Hudson D. J.), подгонка кусочно-линейных кривых по методу НК 474

Халмош (Halmos P. R.), достаточность 42

Хальд (Hald A.), урезанное нормальное распределение 704

Хаммерсли (Hammersley J. M.), оценивание параметров, подчиненных ограничениям (упражнения 18.21—18.22) 104—105

Характеристические функции (characteristic functions), и полнота 256; условные 425—427

Харкнесс (Harkness W. L.), мощность критерия в таблицах 2×2 743

Харли (Harley B. I.), оценивание функции от нормального коэффициента корреляции 393

Хартер (Harter H. L.), кратчайшие доверительные интервалы 162; отношение нормальных размахов (упражнение 23.14) 297; таблицы нормальных меток 673; цензурированные нормальные выборки 704; цензурированные экспоненциальные выборки 706; цензурированные логнормальные выборки 707; таблицы стьюдентизованных размахов 710

Хартли (Hartley H. O.), графики функции мощности для линейной гипотезы 340; доверительные области для нелинейной регрессии (упражнение 28.3) 493; итерационное решение уравнений МП для неполных данных 702; таблицы стьюдентизованного размаха 710

Хатчинсон (Hutchinson D. W.), таблицы биномиальных и пуассоновских доверительных интервалов 165

Хейдер (Hader R. J.), цензурированные экспоненциальные выборки 705

Хейес (Hayes J. G.), обращение матрицы 114

Хейес (Hayes S. P.), номограмма и таблицы для тетрахорической корреляции 410

Хёйланд (Hoyland A.), устойчивое оценивание 673

Хейнэм (Haynam G. E.), мощность критерия Вилкоксона 667

Хемелрейк (Hemelrijk J.), мощность критерия знаков 688 (сноска)

Хеннан Э. (Hannan E. J.), АОЭ для нецентральных χ^2 величин 369; регрессоры 461

Хеннан (Hannan J. F.), мультиномиальные бисериальные методы (упражнение 26.12) 420, мощность критерия в таблицах 2×2 743

Хили (Healy M. J. R.), сравнение линейных моделей (упражнение 28.22) 498

Хили (Healy W. C., Jr.), двойная выборка 826

Хилл (Hill B. M.), оценивание пропорций в смесях распределений (упражнение 17.2) 53; МП-оценка в логнормальном распределении (упражнение 18.23) 105

Хогбен (Hogben D.), моменты нецентральной t-величины 342

Хогг (Hogg R. V.), полнота достаточных статистик 257; полнота, достаточность и независимость (упражнения 23.8—23.9) 296; критерии для k выборок (упражнение 24.6) 347, (упражнение 24.13) 349, 679; критерии ОП, когда область изменения зависит от параметра 317, (упражнения 24.8—24.9) 347; достаточность в экспоненциальном семействе, когда область изменения зависит от параметров (упражнение 24.17) 349

Ходжес (Hodges J. L., Jr.), эффективность критериев 355; таблицы критерия Вилкоксона 662; АОЭ критерия Вилкоксона 667; мощность критерия Вилкоксона 667; устойчивое оценивание 673; АОЭ критерия, основанного на нормальных метках и критерия Вилкоксона (упражнение 31.23)

686; минимальная АОЭ критерия знаков (упражнение 32.1) 712
Холдейн (Haldane J. B. S.), семиинварианты МП-оценки 72; среднее модифицированной МХК-оценки (упражнение 19.14) 138; стандартная ошибка величины X^2 в таблицах $r \times c$ 751, (упражнение 33.9) 783; аппроксимация величины X^2 для таблицы $2 \times c$ (упражнения 33.21—33.22) 787; обратный выбор 795
Холл (Hall W. J.), инвариантность и достаточность 344
Холтон (Halton J. H.), точный критерий независимости в таблицах $r \times c$ 749
Хора (Hora R. B.), фидуциальные выводы 212
Хотеллинг (Hotelling H.), оценивание функций от нормального коэффициента корреляции 393; дисперсия множественного коэффициента корреляции 456; доверительная область для линии регрессии 487, 491—492, (упражнение 28.13) 495; сравнение линейных моделей (упражнение 28.22) 498; устойчивость 627; эффективность критерия ранговой корреляции 646
Хоул (Hoel P. G.), доверительные области для линий регрессии 487, 491, 493, (упражнение 28.13) 495
Хубер (Huber P. J.), устойчивость оценок параметра сдвига 627
Хузурбазар (Huzurbazar V. S.), единственность МП-оценок 60—61, 80; состоятельность МП-оценок 65; доверительные интервалы, когда промежуток зависит от параметра (упражнения 20.13—20.16) 180—181
Хукер (Hooker R. H), данные о погоде и урожаях (пример 27.1) 439
Хьюэтт (Huyett M. J.), цензурированные гамма-выборки 706
Хэджнэл (Hajnal J.), последовательный t-критерий, основанный на двойной выборке 819
Хэмдэн (Hamdan M. A.), выбор групп для критерия X^2 587; полихорическое оценивание 752
Хюрениус (Hyrenius H.), t-величина Стьюдента в составных выборках 625

Цайглер (Zeigler R. K.), двойная выборка 826
Цензурированные выборки (censored samples) 699—707; см. также под названиями распределений
Центр расположения (centre of location) 96
Центральные и нецентральные доверительные интервалы (central and non-central confidence intervals) 144—145

Чанда (Chanda K. C.), критерий Вилкоксона в дискретных случаях 667
Чандлер (Chandler K. N.), соотношения между остатками (упражнение 27.5) 457
Чандра Секар (Chandra Sekar C.), выпадающие наблюдения 709
Частная корреляция и регрессия (partial correlation and regression) 423—444; частная корреляция 423—424; частная линейная регрессия 429—430; соотношения между коэффициентами разных порядков 431—434, (упражнения 27.1—27.3, 27.5) 457; приближенная частная линейная регрессия 434—435; оценивание коэффициентов совокупности 435—437; геометрическая интерпретация 437—439; вычисления (примеры 27.1—27.2) 439—443; выборочные распределения 443—444
Частная связь (partial association) 725—730, 776—781
Чепмэн (Chapman D. G.), граница для дисперсии оценки 34; оценивание стандартного отклонения нормального распределения (упражнение 17.6) 54; урезанное гамма-распределение 706; двойная выборка 826
Чернов (Chernoff H.), асимптотические разложения в задаче о двух средних 202; уменьшение размера критических областей 239 (сноска); измерение эффективности критерия 370, (упражнения 25.5—25.6) 371; распределение критерия согласия X^2 570, 573; АОЭ критерия c_1 (Фишера — Иэйтса) 668, 672; эффективное оценивание с помощью порядковых статистик 703
Чипмэн (Chipman J. S.), теория НК в случае вырождения 124
Чоу (Chow Y. S.), моменты при последовательном оценивании 820
Чупров, коэффициент сопряженности 782

Шапиро (Shapiro S. S.), проверка нормальности 618
Шах Б. (Shah B. K.), логистические таблицы (упражнение 32.12) 714
Шах (Shah S. M.), урезанные биномиальные распределения 707
Шентон (Shenton L. R.), семиинварианты МП-оценки 75; эффективность метода моментов (упражнение 18.28) 107
Шеффе (Scheffé H.), задача о двух средних 192—199, (пример 23.10) 268; линейная функция от χ^2-величин (упражнение 21.8) 217; полнота 254; полнота линейного экспоненциального семейства 256; минимальная достаточность 261; полнота и подобные области 263; несмещенные критерии 277; РНМН критерии для экспоненциального семейства 278; отсутствие полноты у распределения Коши (упражнение 23.5) 295; минимальная достаточность в случае биномиального и равномерного распределений (упражнения 23.11—23.12) 296; особый РНМН критерий для нормального распределения (упражнение 23.16) 297; несмещенный доверительный интервал для отношения нормальных дисперсий (упражнение 23.18) 298; РНМН критерии для пуассоновского и биномиального распределений (упражнения 23.21—23.22) 298; контролируемые переменные 545, 550; полнота порядковых статистик 632, сноска; доверительные интервалы для квантилей 694; толерантные области 698
Шнейдерман (Schneiderman M. A.), последовательный t-критерий 819
Шойер (Scheuer E. M.), таблицы нецентрального t-распределения 342
Шоррок (Shorrock R.), достижимость МГД 31
Шоу (Shaw G. B.), о корреляции и причинности 374—375
Шривастава (Shrivastava M. P.), оценивание функциональной зависимости 538

Эдвардс (Edwards A. W. F.), связь и перекрестное отношение 733

Эзекиел (Ezekiel M. J. B.), доверительные интервалы для многомерной нормальной множественной корреляции 457
Эйзенхарт (Eisenhart C.), предельная функция мощности критерия согласия X^2 582; таблицы критерия серий (упражнение 30.8) 619
Эймос (Amos D. E.), нецентральное t-распределение 342
Эйткин (Aitken A. C.), МГД 24; минимизация обобщенной дисперсии 118; наиболее общая линейная модель (упражнения 19.2, 19.5) 136; ортогональные полиномы 480
Экспоненциальное распределение (exponential distribution), достаточность наименьшего наблюдения для оценивания нижней границы (пример 17.19) 51; достаточность, когда верхняя граница зависит от масштабного параметра (пример 17.20) 52; несмещенная МД-оценка (упражнение 17.24) 56—57; МП-оценка и достаточность, когда обе границы зависят от параметров (упражнение 18.5) 101; поправка на группировку МП-оценки масштабного параметра (упражнение 18.25) 106; порядковые НК-оценки параметров сдвига и масштаба (упражнения 19.11—19.12) 137—138; доверительные границы для параметра сдвига (упражнение 20.16) 181; РНМ критерии для параметра сдвига (пример 22.6) 230; удовлетворяет условию для двусторонней НКО 235; РНМ критерий при отсутствии одномерной достаточной статистики (примеры 22.9—22.10) 238, 240, (упражнение 22.11) 248; РНМ подобные односторонние критерии в случае сложной гипотезы H_0 для масштабного параметра (пример 23.9) 267; независимость двух статистик (упражнение 23.10) 268; неполнота и подобные области (упражнение 23.15) 297; РНМН критерий для сложной гипотезы H_0 о масштабном параметре (упражнение 23.24) 299; РНМ подобный критерий для сложной гипотезы H_0 о параметре сдвига (упражнение 23.25) 299; критерии ОП 329, (упражнения 24.11—24.13, 24.16, 24.18) 348—349; критерий согласия для случайных отклонений (примеры 30.2—30.4) 576, 582, 587, (упражнение 30.6) 619; использование меток 672; урезание и цензурирование 705—706, (упражнения 32.16—32.18) 715—716; обратный выбор 795; последовательный критерий для масштабного параметра (упражнение 34.14) 830
Экспоненциальное семейство распределений (exponential family of distributions) 28; достижимость МГД 31; э. с. р. как характерный вид распределения, допускающего достаточные статистики 46, 48; достаточные статистики, распределение которых имеет тот же вид (упражнение 17.14) 55; э. с. р. и полнота 255—256; РНМН критерии для них 278—292; э. с. р. с областью изменения, зависящей от параметров (упражнение 24.17) 349
Экстремальных значений распределение (extreme-value distribution), МП-оценка (упражнение 18.6) 101; цензурированные данные 706—707
Эллисон (Ellison B. E.), толерантные интервалы 178
Эль-Бадри (El-Badry M. A.), МП-оценка групповых вероятностей в таблице $r \times c$ (упражнение 33.29) 789
Эммон (Ammon O.), данные о цвете глаз и волос (упражнение 33.3) 781
Эндрюс (Andrews F. C.), АОЭ критерия Крускала — Уоллиса 676
Энскомб (Anscombe F. J.), МП-оценка в отрицательно-биномиальном распределении (упражнения 18.26—18.27) 106; выпадающие наблюдения 710; таблицы для последовательной проверки качественных признаков 808; последовательное оценивание 821; последовательный критерий для экспоненциального распределения (упражнение 34.14) 830
Эпстейн (Epstein B.), свойство независимости экспоненциального распределения (упражнение 23.10) 296; цензурирование в экспоненциальных выборках 705, (упражнение 32.17) 716
Эффективность (efficiency) оценивания, и корреляция между оценками 35—36; определение 36—37; измерение 37—38; разбиение ошибки в случае неэффективной оценки (упражнение 17.11) 55; эффективность МП-оценок 67—72, 83—86; эффективность метода моментов 97—100, (упражнения 18.9, 18.17, 18.20, 18.27—18.28) 102—104, 106—107; э. о. и мощность критериев 231—233, (упражнение 22.7) 233; э. о. и АОЭ критериев 367
Эффективность критериев (efficiency of tests) 351—353; см. Асимптотическая относительная эффективность

Юл (Yule G. U.), развитие корреляционной теории 374—375; обозначения частной корреляции 423; асимптотические распределения частных коэффициентов 443; связи между коэффициентами различных порядков (упражнение 27.1) 457; данные о прививках (пример 33.1) 722; коэффициенты связи и коллигации 723; принцип инвариантности для мер связи 732

Янг (Young A. W.), стандартная ошибка величины X^2 в таблицах $r \times c$ 750

c_1 (двухвыборочный) критерий (c_1 (two-sample) test) 667—675
F-распределение (F-distribution) нецентральное 337—341
H_0, проверяемая гипотеза, нулевая гипотеза (H_0, hypothesis tested, null hypothesis) 222
H_1, альтернативная гипотеза (H_1, alternative hypothesis) 222
k-выборочные критерии (k-sample tests), (примеры 24.4, 24.6) 314, 328; (упражнения 24.4—24.7) 346—347; (упражнения 24.11—24.13) 348—349; 675—679
X^2, критерий согласия (X^2 test of fit) 561—593, его разбиения 601—602; X^2 в таблицах 2×2 736, 744; X^2 в таблицах $r \times c$ 745—751, (упражнение 33.9) 783; разбиения X^2 в таблицах $r \times c$ 769—774; дисперсионные критерии 774—776, (упражнения 33.20—33.22) 787; X^2 в таблицах с несколькими входами 776—781, (упражнения 33.7—33.8) 782—783; см. Критерии согласия
χ^2-распределение (χ^2-distribution), см. Гамма-распределение, Нецентральное χ^2-распределение.

М. Кендалл, А. Стьюарт

Статистические выводы и связи

М., 1973 г., 900 стр. с илл.

Редактор *Д. М. Чибисов*

Техн. редактор *И. Ш. Аксельрод*

Корректоры *Т. С. Плетнева, Л. С. Сомова*

Сдано в набор 18/VII 1973 г. Подписано к печати 12/XII 1973 г.
Бумага 60×90¹/₁₆, тип. № 1. Физ. печ. л. 56,25. Условн. печ. л. 56,25.
Уч.-изд. л. 56,94. Тираж 14 500 экз. Цена книги 4 р. 34 к.
Заказ № 731

Издательство «Наука»
Главная редакция физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

Ордена Трудового Красного Знамени
Ленинградская типография № 2 имени Евгении Соколовой
Союзполиграфпрома при Государственном комитете Совета
Министров СССР по делам издательств, полиграфии
и книжной торговли
198052, Ленинград, Л-52, Измайловский проспект, 29

СТАТИСТИЧЕСКИЕ ВЫВОДЫ И СВЯЗИ



КЕНДАЛЛ
СТЬЮАРТ